# A Literature Review on Application of Sentiment Analysis Using Machine Learning Techniques

**Anvar Shathik J.[1, 2]& Krishna Prasad K.[3]**

[1]Research Scholar, Srinivas University, Mangaluru, Karnataka, India

[2]Assistant Professor, Department of Cloud Technology and Data Science, College of Engineering & Technology, Srinivas University, Mangaluru, Karnataka, India

[3]College of Computer Science and Information Science, Srinivas University, Mangalore, India

Email: anvarshathik@gmail.com

**Subject Area:** Computer Science.
**Type of the Paper:** Review Paper.
**Type of Review:** Peer Reviewed as per |C|O|P|E| guidance.
**Indexed In:** OpenAIRE.
**DOI:** http://doi.org/10.5281/zenodo.3977576.
**Google Scholar Citation:** IJAEML.

**How to Cite this Paper:**

Anvar Shathik, J. &Krishna Prasad, K. (2020). A Literature Review on Application of Sentiment Analysis Using Machine Learning Techniques. *International Journal of Applied Engineering and Management Letters (IJAEML)*, *4*(2), 41-77.
DOI: http://doi.org/10.5281/zenodo.3977576.

**International Journal of Applied Engineering and Management Letters (IJAEML)**
A Refereed International Journal of Srinivas University, India.

© With Authors.

This work is licensed under a Creative Commons Attribution-Non-Commercial 4.0 International License subject to proper citation to the publication source of the work.
**Disclaimer:** The scholarly papers as reviewed and published by the Srinivas Publications (S.P.), India are the views and opinions of their respective authors and are not the views or opinions of the S.P. The S.P. disclaims of any harm or loss caused due to the published content to any party.



Electronic copy available at: https://ssrn.com/abstract=3674982

# A Literature Review on Application of Sentiment Analysis Using Machine Learning Techniques

**Anvar Shathik J.[1, 2]& Krishna Prasad K.[3]**

[1]Research Scholar, Srinivas University, Mangaluru, Karnataka, India

[2]Assistant Professor, Department of Cloud Technology and Data Science, College of Engineering & Technology, Srinivas University, Mangaluru, Karnataka, India

[3]College of Computer Science and Information Science, Srinivas University, Mangalore, India

Email: anvarshathik@gmail.com

## ABSTRACT

Many businesses are using social media networks to deliver different services and connect with clients and collect information about the thoughts and views of individuals. Sentiment analysis is a technique of machine learning that senses polarities such as positive or negative thoughts within the text, full documents, paragraphs, lines, or subsections. Machine Learning (ML) is a multidisciplinary field, a mixture of statistics and computer science algorithms that are commonly used in predictive and classification analyses. This paper presents the common techniques of analyzing sentiment from a machine learning perspective. In light of this, this literature review explores and discusses the idea of Sentiment analysis by undertaking a systematic review and assessment of corporate and community white papers, scientific research articles, journals, and reports. The goal and primary objectives of this article are to analytically categorize and analyze the prevalent research techniques and implementations of Machine Learning techniques to Sentiment Analysis on various applications. The limitation of this analysis is that by excluding the hardware and the theoretical exposure pertinent to the subject, the main emphasis is on the application side alone. The limitation of this study is that the major focus is on the application side thereby excluding the hardware and theoretical aspects related to the subject. Finally, this paper includes a research proposal for e-commerce environment towards sentiment analysis applying machine learning algorithms.

**Keywords:** Machine Learning Techniques, Sentiment Classification, Sentiment Analysis Applications.

## 1. INTRODUCTION :

Social networking channels like Twitter, Facebook, Instagram, and WhatsApp have stormy contact environments, it is imperative to relay sensitive knowledge about people's opinions, moods and feelings on any product, concept or policy through these social network channels(Yi & Liu, 2020)[1]. To both customers and suppliers, this data is valuable. During any online shopping, consumers usually check other people's opinions about the product. Based on the customer's sentiment, the manufacturer can learn about its product benefits and drawbacks. Although both business organizations and individuals can get profit from these opinions, the sheer number of these opinions on text data is daunting for the users. For researchers, it is a very interesting area to examine and sum up the opinions conveyed in this broad opinion text content. This modern area of study is also known as Sentiment Analysis or Opinion Mining(Vohra & Teraiya, 2013)[2].

In the period of machine learning, machines are left to think and solve the problems by finding the patterns in every data set on their own. Examination of secret trends and patterns helps predict and avoid potential problems. A machine-learning algorithm uses a specific type of data to reply to more questions using the patterns hidden in that data. The importance of machine learning has now been recognized by many companies dealing with large quantities of data. Besides, cost-effective computational processing and data-storage options have allowed the creation of models that analyze large volumes of complex data quickly and precisely. To obtain the highest value from big data, businesses need to know precisely how



Electronic copy available at: https://ssrn.com/abstract=3674982

to fit the correct algorithm with a particular learning process or resources (Machine Learning & its Applications Outsource to India, 2020)[3].

Thisis highlighted with these following subsections: i) Need of sentimental analysis, ii) Sentiment Classification, iii) Approaches for sentiment analysis, iv) Sentiment analysis research fields, v) Machine learning based sentiment analysis methodology, vi) Architecture of sentimental analysis for social media anal, vii) Outline of machine learning techniques, viii) Applications of sentimental analysis using machine learning techniques, ix)Use cases of sentimental analysis, x) Summary of related work xi) Discussion, xii)Research Gap, xiii)Research Agenda and analysis, xiv) Research proposal, xv)Strength Limitation Opportunity Challenges (SLOC) analysis, xvi)Implementation suggestion, xvii)Conclusion.

## 2. RESEARCH GOAL AND METHOD :

The literature review is performed to recognize the applications and solutions of sentimental analysis for the analyzing and classification of views using Machine Learning (ML) approaches. These techniques, therefore, need to find various solutions to the problems found. It also requires certain classifications and methods to evaluate the solutions to certain limitations. Finally, the review ends with the stability and significance of the proof. We focused on the following research questions:

- RQ1: What are sentimental analysis applications used Machine learning techniques?
- RQ2: How much activity was carried out recently?
- RQ3: What are the different classification, methodology and techniques that were used?
- RQ4: How do the different techniques prove themselves as best?

## 3. OVERVIEW OF SENTIMENTAL ANALYSIS :

The method of collecting primary information from unstructured and unoriented textual materials from various social media and website resources, such as chatting on social networks like Twitter, WhatsApp, Facebook, live blogs, or comments can be described as sentimental analysis.(Yi & Liu, 2020)[1].The process for examining and sum up the opinions expressed in these enormous opinions generated by users is commonly named as opinion mining (Vohra & Teraiya, 2013) [2].

### 3.1 NEED OF SENTIMENTAL ANALYSIS

We are living in a "data age" today. As the number of users on social media sites like Twitter grows rapidly, numerous opportunities and new scope have been opened up to businesses trying to keep track of consumer feedback and product opinions. Tweets are a useful source of opinion or sentiment for businesses, governmental bodies, or individuals in the sense of Twitter's social network (products, people, trends, events, etc.). Nonetheless, Twitter generates a very large number of tweets every day (21 million tweets every hour, as reported in 2015). Therefore the method of sentiment analysis needs to be automated to promote the tasks of evaluating the views of the public without the need for millions of tweets to be read manually(Jain & Dandannavar, 2016)[4].

### 3.2 SENTIMENT CLASSIFICATION

Sentiment analysis is an automated method of determining whether a usage-produced text conveys a positive, negative or common view of an object (i.e. the item, the individual, the subject, the case, etc.). Sentiment classification can be achieved at the four levels such as Document level, Sentence level, and Aspect or Feature level(Vohra & Teraiya, 2013) [2].

A. **Document-level:** The Document-level uses the entire documents to categorize it into a positive or negative class as a simple information category.
B. **Sentence level:** In the Sentence level, the sentiment classification categorize any sentence as subjective or objective, and then it categorize into a positive, negative, or common class.
C. **Aspect or Feature level:** This type of sentiment classification discusses the identification and extraction of item features from source data.

### 3.3 APPROACHES FOR SENTIMENT ANALYSIS

There are several techniques available for sentiment analysis:

A. **Lexicon based approach:** A dictionary containing both positive and negative terms used by Lexicon is applied to assess the polarity of opinion. The count of optimistic and pessimistic words is discussed in the text. If the text is more positive, a positive score will be assigned to the text. The text is awarded a negative score if it has high amount of negative or pessimistic words. If the text contains the same number of good and bad terms, a neutral score is given. A lexicon of opinion (positive and negative



Electronic copy available at: https://ssrn.com/abstract=3674982

opinions) is developed to finalize the word is positive or negative. There are numerous ways to build and compile a dictionary (Medhat et al., 2014)[5].

- **Dictionary-based approach:** A small number of words of opinion with established guidelines are gathered manually(Medhat et al., 2014)[5]. In corpora like WordNet or thesaurus, synonyms and opposite from these words are then searched and appended to the group. The collection decreases slowly until there are no new terms. This method has the inconvenience of depending on the dictionary scale, the intensity of the sentiment classification. As the dictionary size increases, this approach is wrong(Jain & Dandannavar, 2016)[4].
- **Corpus-based approach:** They rely on massive corporations for syntactic and semantic opinion patterns. The created words are context-specific and it needs a large dataset labelled (Jain & Dandannavar, 2016)[4].

B. **Machine learning-based approach:** Machine learning techniques in the classification of sentiment depends on the use of well-known machine learning technology on text data. The classification of the sentiment based on machine learning can be categorized primarily into supervised and unsupervised methods of learning(Aydogan & Akcayol, 2016)[6].

- **Supervised learning:** Supervised methods of learning rely on labeled training manuals. Supervised learning is an effective classification method and has been used with very promising results for classifying opinions. The regularly used supervised classification techniques in sentiment analysis are Support Vector Machine (SVM), Naïve Bayes (NB) Maximum Entropy (ME), and Artificial Neural Network (NN) and Decision Tree (DT) classifiers. Some less commonly used algorithms are Logistic Regression (LR), K-Nearest Neighbor (KNN),Random Forest (RF),and Bayesian Network (BN)(Aydogan & Akcayol, 2016)[6].
- **Unsupervised learning*:*** This technique does not use pre-listed data to train the classifier, unlike supervised learning. The more common instance of unsupervised machine learning algorithms are K-means and Apriori Algorithms. Unsupervised machine learning may also be divided into clusters and associations(Ahmad et al., 2017)[7].

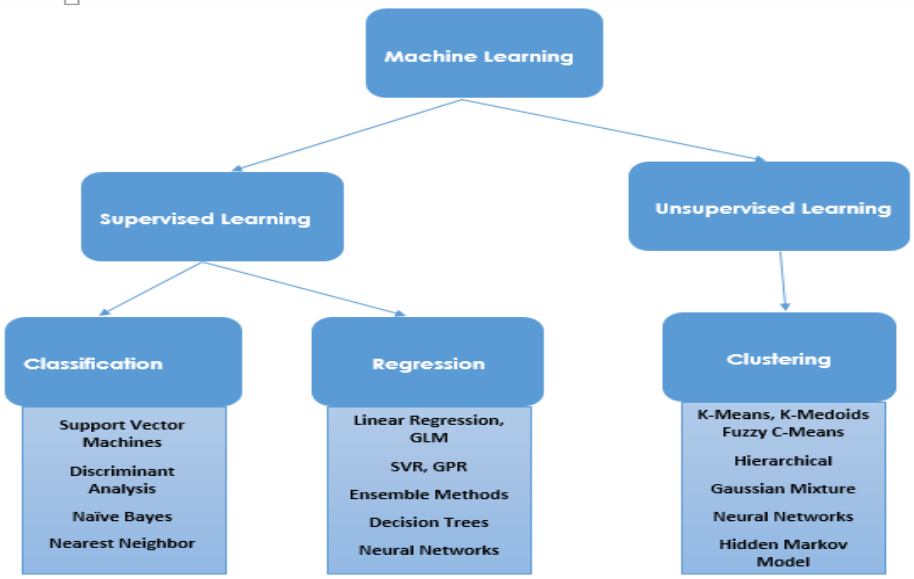


**Fig. 1**: Machine learning based approach(Medhat et al., 2014)[5]

C. **Hybrid based approach:** The hybrid-based approach uses both ML and lexicon-based classification approach. Few research techniques propose a mixture of lexicon-based and automated learning techniques to enhance the classification of sentiment. This hybrid approach is primarily advantageous as it can achieve the best of both. The combination of Lexicon and Learning has demonstrated increased accuracy(Jain & Dandannavar, 2016)[4].

### 3.4 SENTIMENT ANALYSIS RESEARCH FIELDS

Subjectivity identification, opinion divination, aspect-based sentiment summary, text overview, contrast point overview, product feature derivation and the detection of opinion spam bots are key areas of



Electronic copy available at: https://ssrn.com/abstract=3674982

research in sentiment analysis(Vohra & Teraiya, 2013)[2].
**Subjectivity Detection:** It is to examine whether the text is expressed or not.
**Sentiment Prediction:**It is to predict whether the text is positive or negative.
**Aspect Based Sentiment Summarization:** It provides an opinion outline of feelings in the form of high ratings or credits of product characteristics.
**Text Summarization:** It creates a few phrases that sum up a product's ratings.
**Contrastive Viewpoint Summarization:** It provides the summary of Contrastive Viewpoint emphasizes contradictory views.
**Product Feature Extraction:** It is basically a job that derives from its review the product features.
**Detecting opinion spam:** It provides the identification of fraudulent or false opinions from feedback that requires the detection of spam opinion.

### 3.5 LITERATUREREVIEW

In general opinion research at the starting of the 20th Century, the science of sentiment analysis and opinion mining has a strong basis. When online product reviews were required and accessible in the middle of 2000, they finally became a major research subject. Just 101 articles on this subject were published in 2005, while almost 5,699 were published in 2015. This means that over a decade sentiment analysis has increased almost 50 times, making it one of the most quickly expanding fields of study in previous years(Mäntylä et al., 2018)[8].Throughout the early days of the internet, a person was able to seek feedback from his friends, neighbors and relatives before taking any decision. Opinion sampling, surveys, and general public opinion on its products or services were conducted by organizations. As the World Wide Web has come and particularly with the production and adoration of Web 2.0, where the focus on content generated by users has changed significantly the way the individual expresses his opinion or views. Now people can offer their thoughts, opinions, feelings, blogs, social platforms, forums, and reviews on their own personal web pages. Thanks to rich and diverse data generated in Web 2.0 applications, the field of opinion mining has advanced quickly(A. Kumar & Sebastian, 2012)[9].Research into the shifts in the subjects found that social networking such as Twitter and Facebook are more focused on the most recent articles from the year 2014 to 2016. In recent year's mobile devices, stock markets, and human emotions were other topics that have become popular(Mäntylä et al., 2018)[8].

In2012, the time of major information advances, and in 2013, the Big Data Analytics region is becoming more popular. Big data is due to advanced techniques for data processing with massive and high-dimensional data, dramatically increased storage capabilities, and complex data formats. In this area, the Big Data requires state-of-the-art technology and/or techniques to solve the different computational times to collect useful data without sensitive data loss. A new and rapidly expanding field of research has recently been proposed: machine learning to overcome these problems. Master learning algorithms have generally been thought to learn from large volumes of data and to find useful and valuable information (Swathi &Seshadri, 2017) [10]. Brief information about different types of algorithms used for sentiment analysis is given. Sentimental analysis is defined as the analysis of opinions, thoughts, sentiments, and subjectivity of text are given. Recently introduced algorithms, sentimental analysis techniques are discussed, and also the importance of some fields such as transfer learning, feelings detection, and constructing resources are discussed. The main purpose of this survey is the categorization of recent articles, 54 of the latest published articles which are based on sentiment analysis were categorized and summarized (D.Kawade & Oza, 2017) [11]. Social media acts as an important source where one can interact and can be able to fulfill their demands. This brings both satisfaction for customers and also companies. The traditional-based analysis is difficult to analyze, there are some challenges to overcome this problem. Some methods for analyzing feelings, such as prediction of user subjects, polarity of feelings scores, qualitative analysis and a large data mining application, cross-domain classification of feelings, identification of emotional differences, meaning and theme detection, classification of hashtag sentiment rates, sales forecasts, etc. are used. It also addressed briefly the complexities of sentimental analytics to do the job. Some of the challenges such as parallel computing for massive data, sarcasm, grammatically incorrect words, review the author's segmentation, handling noise, and dynamism. (Patil & Atique, 2015)[12].The consumer can compare products according to the people's reviews on these products. So, for making this more successful they have produced supervised techniques for the consumer reviews.There are two types of methods are mentioned that is, association rules techniques and



Electronic copy available at: https://ssrn.com/abstract=3674982

naïve Bayes classifiers to categorize the features of the products that according to the needs of consumers. This analysis is not only based on the ratings, but the important character also is, and this sentimental analysis compares and identifies the preferred products which make comfortable for the consumer. An empirical evaluation, they have mentioned two classifiers namely, Naïve Bayes classifiers and class association rules (Yang et al., 2010) [13]. The reviews of sentiments are classified accurately by the algorithms of machine learning such as bag-of-words, n-gram, naive Bayes classifier, and natural language processing. Then the user's sentiments are categorized as positive, neutral, negative, the top features of the product will make the customer attract towards that particular product. This work also says that the future scope of reviewing the products are based on opinions in several languages, copying drawback of mapping slangs, copying with mocking opinions, and then providing comparative opinion between two products for one best and copying with anaphora resolution(Gopu & Swarnalatha, 2017)[14]. The sentiment analysis uses natural language processing to naturally classify and derive the emotion from the text and as a result, it has a variety of applications in the consumer sector, for example. Transfer education has also emerged as a new method of machine learning that utilizes existing knowledge to solve issues and to generate forecast results. It also contains the prospect of sentimental analysis such as the appliance of cross-domain shift learning aspects which has not been fully explored, then solving negative problems of text data by using transfer learning becomes very difficult. And they conclude that in future Aspect level sentiment analysis for small texts is considered as the most promising research technique. (R. Liu et al., 2019) [15]. Machine learning algorithms like Naive Bayes, Support vector machine, and Maximum entropy classifier algorithms can be used on the sentimental analysis of huge data. Using these techniques, a huge volume of data can be utilized to get optimized and strategical decision-making capability. sentimental analysis is also called opinion mining which gives a brilliant and human-like brilliance which analyzes and respond emotions, the user show in social media like Facebook, Yammer, Twitter, microblogs which provide a tremendous amount of data every day in textual or numerical forms and these are classified as structured, semi-structured and non-structured and then later they are categorized as positive, negative and common based on user's attitude towards a particular topic for analysis purpose(Naiknaware et al., 2017)[16]. Sentiment analysis helps to do a review of the movies, product, and customer opinion on products. The role of sentiment analysis in natural language processing is to remove positive or negative polarities from social media messages. Digital social networks are growing increasingly, and culture focused on online media has affected young scientists in their research in the study of opinion. Organizations that are actually keen to determine their clients or the public opinion on their social media goods. Internet services should be able to test social media data on blogs, web forums, articles, tweets and user feedback (J. Singh et al., 2017) [17].

### 3.6 MACHINE LEARNING BASED SENTIMENT ANALYSIS METHODOLOGY

The polarity of an analysis data is calculated by various techniques. Machine learning Basic sentiment analysis technique is the most popular and efficient. As discussed below, the polarity in analysis data and the most successful algorithm are calculated(Yogi & Paudel, 2020)[18].

**Data collection**

For any kind of text classification task-specific in size as to the number of words, data sets can be used. Such data sets were used after slight preprocessing for sentiment analysis such as case folding, word deletion, etc.

**Data Preprocessing**

This pre-processing phase seeks to prepare text data for further processing.

**Feature Selection and Feature Vector Construction**

A computer is not able to process text data straight away, which is an inherent problem. Text data must also be numerically interpreted. Terms are usually used as the characteristics of the text. This gives the text representation a high dimension. Features need to be filtered to reduce dimensions and remove noise to improve classification performance and processing efficiency.

**Classification Algorithms for Sentiment Analysis**

Several popular and commonly used classification algorithms such as the Multinomial Naïve Bayes Algorithm or the K-Nearest Neighbor Algorithm are commonly used to identify sentiment polarity of users' opinions based on given opinion data Support Vector Machines Algorithm.

**Evaluation Metrics**

Measuring every algorithm's output using parameters such as confusion matrix, efficiency, recall, and F-



Electronic copy available at: https://ssrn.com/abstract=3674982

measurement

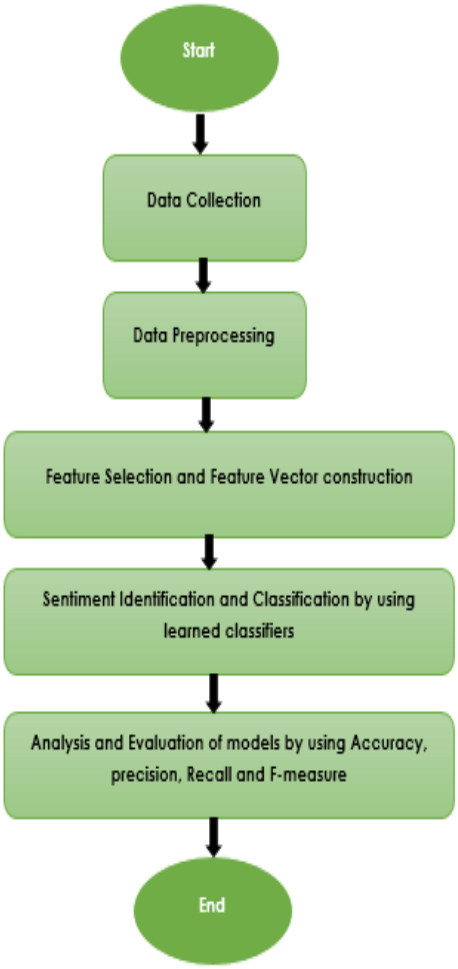


**Fig.2**: Flow Chart of machine learning based sentiment analysis technique(Yogi & Paudel, 2020.)[18].

### 3.7 ARCHITECTURE OF SENTIMENTAL ANALYSIS FOR SOCIAL MEDIA ANALYTICS

The area of feelings investigation which investigates feedback, customer opinion, feeling assessment, writing emotions, and attitudes. The review of the product as many is negative and positive is a decision-making process. The key method of the rejection norm and the classification of negative and positive feelings earned by the users or customers in the social community(Mahendran & Mekala, 2018)[19].

**Data Pre-processing:** The preprocessing technique is more useful to identify and remove meaningless, noisy, and incompatible data.

**Eliminating URLs:** URLs will not help to inspect the emotion in the non-formal text.

**Questions:** The terms question will not help to reduce the ambiguity of polarity such as when, when, who, how, etc.

**Removing Special Characters:** Special characteristics like, and, (), [] {}/' are separated to eliminate inconsistencies by the function of polarity.

**Removal of Retweets:** The re-tweeting process doubles the tweet of another user and redistributes it to peers. This also happens whenever a user tries to tweet another user. Retweets are normally reduced.



Electronic copy available at: https://ssrn.com/abstract=3674982

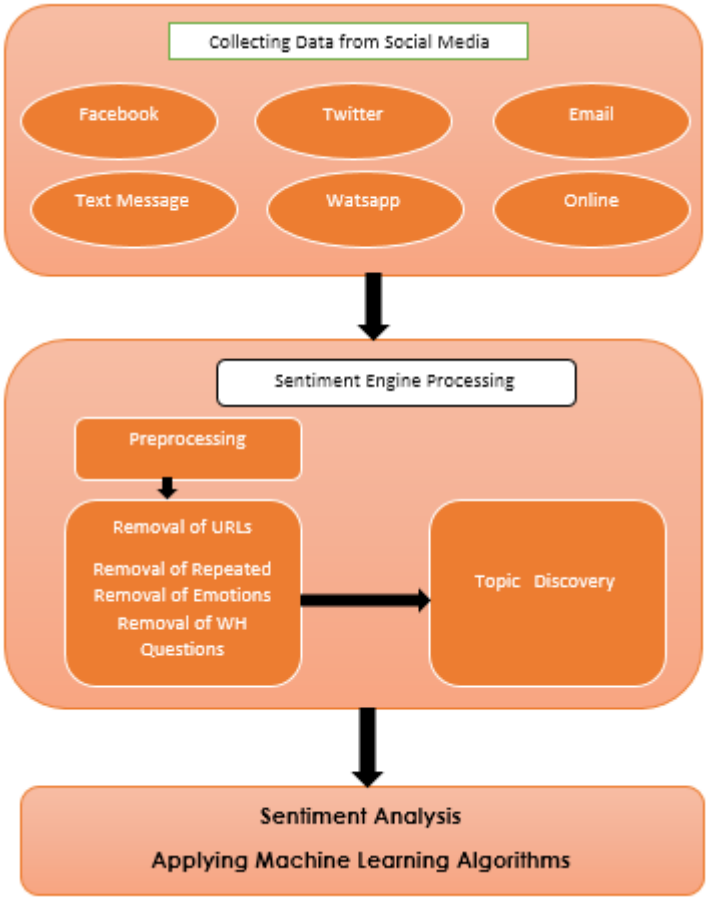


**Fig. 3**: Architecture of sentimental analysis(Mahendran & Mekala, 2018)[19].

### 3.8 OUTLINE OF MACHINE LEARNING TECHNIQUES

Mechanical learning is an Artificial Intelligence (AI) branch, which investigates machines for the establishment of new knowledge and skills and the identification of existing knowledge. In the area of data mining, computer vision, processing of natural languages, search engines, biometrics, medical diagnostics, credit card fraud detection, a market analysis of stocks, DNA sequence, speech and handwriting recognition, strategy games and robotics, machine learner has been widely used(Patel, 2018)[20]. The popular machine learning algorithms are

**1. Linear Regression:** A linear regression is defined as thevalue of the dependent or reliant variable is estimated using independent variables for statistical techniques? A relationship consists of mapping a dependent and independent variable on a line and that line is known as regression line shown by Y= a*X + where Y is the Dependent variable, X is the Independent Variable is the Intercept and a is the slope.

**2. Logistic Regression:** This approach is used to define the discrete dependent variable from the set of separate variables. Logistic regression provides the coefficients to estimate a probability logistic transformation.

**3. Decision Tree:** For classification and regression, the decision tree may be used both with a structure like a tree. The best attribute of a dataset is put in a decision tree building algorithm, then the training data set is divided into subsets. Decision trees are designed to generate a training model to predict the class or value of the destination variable.

**4. Support vector machine:** A Binary Classifier (BC) is a Support Vector Machine (SVM). On the n-dimensional point, row data is drawn. In this, a hyper plane separating the data sets is drawn. This enhanced separation maximizes the training data margin.

**5. Naive-Bayes:** This method is based on the theorem of Bayes used by increasingly sophisticated classification methods. It is a classification technique. It learns how an entity with certain characteristics belonging to a certain category or class is possible.



Electronic copy available at: https://ssrn.com/abstract=3674982

**6. KNN:** This technique is used for classification and regression. This is one of the simple machine learning algorithms. It saves the cases and searches most k-neighbors it resembles for new data. It saves the cases. With a testing dataset, KNN makes clear predictions.
**7. K-means Clustering:** It is an unsupervised algorithm for learning to reach the cap. The initial partition is achieved by Euclidean distance for grouping the datasets into clusters.
**8. Random Forest:** It is the category of the supervised algorithm. In a random forest algorithm, i.e. set of many classification trees, is generated by many decision trees taken together. This can be used for both regression and classification. The Decision Tree algorithm has rules for the given training data set with targets and features.
**9. Dimensionality Reduction Algorithms:** This implies that the number of random variables is decreased by acquiring those key variables. Function extraction and selection of features are methods for reducing dimensionality. The main component analysis can be done by Principal Component Analysis (PCA), which is a method of removing primary variables from a wide range of variables.
**10. Gradient boosting and Ada Boost Algorithms:** The algorithm of gradient boosts is a classification and regression algorithm. AdaBoost selects only those features that enhance model prediction. It operates by selecting a base algorithm such as decision trees and refining it iteratively by taking into account the wrong examples in the training data set.

## 3.9 APPLICATIONS OF SENTIMENTAL ANALYSIS USING ML TECHNIQUES

**Twitter sentiment analysis using Natural Language Processing techniques**

For Opinion mining, Suryawanshi et al., (2020) [21] says that a large quantity of data is needed from people's thoughts on Twitter. In natural language processing, several approaches help retrieve tweets from Twitter directly. The tweets are unstructured. It needs to process and clean tweets to achieve opinion mining, structured data. Until the study, data are processed by the elimination of all links, hashtags, and capitalized terms, repeated phrases, and short-form terms, errors in spelling, special symbols, twitter characters, and remaining text. Data cleaning includes extracting and converting text to the data frame, removing text URLs, deleting stop words such as (the, a, ...), usernames, profiles, deleted numbers, and unused spaces, erase dots and converting Emoji's from Latin to ASCII. The process of removing data involves removing text from tweets. It just have the text of tweets after processing and cleaning. This word for tweets is one by one and provides the word meaning from WordNet using the Vader lexicon tool. A growing word's value is measured and marked as a tweet sentiment score. If opinions are reached, it classifies every tweet as positive, normal, and negative by using a machine learning classification.

**Twitter sentiment analysis using Multinomial Naive Bayes and Logistic Regression:**

Sentiment analyses on Twitter are challengeable. Some of the challenges are (i) some tweets are commonly recognized in unceremonious languages and even a few short words display only small signs of feeling. (ii) Also, twitters are widely used for hashtags, URLs, abbreviations, emoji, and acronyms. Concepts such as tweets before delivery, extraction methods, and table design, use the precision of various machine-learning algorithms. Machine learning methods are utilized to exercise the algorithms to include Multinomial Naive Bay and the logistic regression algorithm with the train data on the test results. The airline sentiment data set and IMDB analysis datasets are the data sets considered by the author. All types attain the good results in machine learning when practicing the Count Vectorizer functionality. The author observes that the test set results from the Logistic Regression with Count Vectorizer features (*Sentamilselvan K*, 2020) [22].

**Twitter Sentiment Classification using Machine Learning Techniques**

The tweet classification is performed applying ML methods such as the Multilayer Propon (MLP), Naïve Bayes, Fuzzy Classification, Decision Tree, and Support Vector Machines (SVM.). Such techniques help to examine various component vectors with a doled-out class so that the connection dependence between the assessment and each element is distinguished. Performance measurements such as accuracy, duration, alert, and F calculation are analyzed in the Twitter dataset. These techniques are evaluated. Accuracy is ranging from 72.66 to 92.34%, Accuracy from 72.16 to 90.12%, Recall is 72.81% to 94.34%, and F-measurement is ranging from 72.4 to 92.3% for classification methods. Results showed that in the Twitter Sentiment review SVM surpassed all others (Godara & Kumar, 2020) [23].

**Sentence Level Sentiment Analysis from News Articles and Blogs using Machine Learning Techniques**



Electronic copy available at: https://ssrn.com/abstract=3674982

The detection and measurement of sentence-level refusals from news articles and forums follow two measures commonly used for pre-processing and posting analytics. For research, the authors used data from the BBC News Report and the multisensory dataset. Methods including data cleaning, number elimination, document conversion to lower case, stroke, and frequency of the text. The recognition of the sentiment is done by the use of opinion and syuzhet packages with a dictionary technique and a lexical methods. Sentiment cluster is better for the sentence level, because with more parameters, such as, positive term, negative word, down the tower, amplifier, de-amplifier, adverse combination, etc., the polarity of every line is determined. The Naïve Bayes (NB) and the Vector Support Machine (SVM) classification systems used are utilized here. Naive Bayes reaches 96.46% and Vector Machine helps algorithms reach 94.16%. Naive Bayes obtain more good results that compare with the SVM to the probability and features of each word in the given sentences(Shirsat et al., 2019)[24].

**Online customer feedback in the restaurant sector of the hospitality industry**

To determine the performances of several supervised machine learning algorithms and conduct secondary research to ascertain whether opinion mining can be extended to find fake positive and negative reviews through data collection, data purification, data transformation, and reduction by means of the string-to-Word Vector. The classification algorithm on a dataset is known as the classification algorithm. Naive Bayes, Decision Tree, Help Vector Machine, K-NN, and K-star are the classifiers used in the analysis. For each of the five supervised machine learning algorithms or classifications applied to both datasets a confusion matrix is generated. The authors conclude from the results that the SVM algorithm has the best precision, the most accurate, the lowest negative detection rate, the lowest down rate, the highest F1 metric, the lowest error rate and the longest time for both the datasets, and k- the closest neighbor(Yadav et al., 2020)[25].

**Sentiment analysis for Indian premier league using a machine learning technique**

Gujar & Pardeshi, (2020)[26] uses the specific machine learning approach to describe an opinion mining method on social media. The tweet list of 2016 Indian premier league tweets that are viewed using Twitter's API (Application Programming Interface) services after the algorithm relies on what the hashtag (# IPTEAM) is. The performance of the Random Forest algorithms is evaluated with the accuracy, precision, and sensitivity of supervised machine learning algorithms already implemented.

**Sentiment analysis for Czech language using a machine learning technique**

In the Czech language, the Opinion mining is difficult due to its very flexible word order and its various word formations. The researchers targeted sentiment analysis in Czech. It exposed and addressed unexpected hurdles, created required evaluation data lines, and even established new innovative approaches for the analysis of sentiments. For the classification of items, a material analysis approach is used that is accompanied by feature extraction and classification steps[27]

**Real-Time Prediction of BITCOIN Price using Machine Learning Techniques and Public Sentiment Analysis**

By machine learning techniques and Opinion mining, decide the predictable price path for bitcoin in the US Dollar. IT studied and tracked the principles of machine learning in Twitter and Reddit posts for extracted tweets, and studied the relation between Bitcoin price fluctuations and tweets. To construct a prediction model and yielda detailed analysis of upcoming market rates, it explored various algorithms for masterminding using supervised learners. It is also quite difficult to make the right predictions because of the difficulty in determining the exact existence of a model of Time Series (ARIMA). Instead, implementing long-term memory cell (LSTM) Recurrent Neural Networks (RNN). Then analyzed bitcoin price forecasts of long-term memory models with more effectiveness and compared the future of the bitcoin cost and feel tests to the standard approach (ARIMA) with the long-term memory (LSTM) techniques. The RMSE of the LSTM is 198.448 (mono feature) and 197.515 (several-feature), while the ARIMA RMSE model is 209.263 that shows the multi-function LSTM yields the most authentic results(Raju & Tarif, 2020)[28].

**Sentiment analysis on e-sports for education curriculum using Naive Bayes and support vector machine**

The analysis was obtained by raking through social media onto Twitter and utilizes an algorithm of Classification (Naïve Bayes) and Support Vector Machine (SMOTE). It uses Data Mining CRISP-DM (Cross Industry Standard Model). A Rapid Miner version 8.2 tool is used to process the text data. Transform Case on Rapid Miner will be used to downgrade all capital terms. Rapid Miner Tokenize is



Electronic copy available at: https://ssrn.com/abstract=3674982

used to erase punctuation marks, and eliminate unwanted signs or special characters from each document in the manner that all special structures are separate structures and called user names. Classify and test the templates, Filters Token by Length and Filter Stop terms. The SMOTE process optimizes the algorithm, the accuracy is 66.92%, accuracy is 61.31% and the recovery value is 92.15%. The algorithm for Naïve Bayes could, therefore help predict the success of e-sports for student curriculum(Ardianto et al., 2020)[29].

**Recognizing Public Sentiment Polarity: Analysis of Smart Phone Industry Using Machine-Learning Approaches**

A comparison is performed in the present study between five major smart smartphone brands that have the market share – Samsung, Apple, Huawei, and Xiaomi and Oppo-based on sentimentally divided prices. Public assessments are collected online and sentiment ratings are measured to create polarity of public sentiments to the popular brands. Analysis of the sentiment is based on essential functions such as tokenization, lemmatization, stopping word recognition, recognition of the named person, and identification of the negative term. For polarity measurements in public comment, the Vader algorithm were used(Imtiaz & Islam, 2020)[30].

**A Machine Learning Approach to Negation and Speculation Detection for Sentiment Analysis**

The most critical role in sentence analysis is the detection of negative knowledge and the detection of speculation. This type of sentimental analysis will be achieved by machine learning. This approach strengthens the description of the text's polarity. The analysis has two stages, for example, the first step: negative signs and uncertainty are found in the document. In the second phase: this signal is calculated in its entirety. The Simon Fraser University Review corpus is responsible for this function, which is used in opinion mining. Some methods, such as rigorous approaches, advanced linguistic analysis tools for automated evaluation of views, significant effort in the extraction of information, detect negative information. Domain identification of subjectivity increases recognition of polarity. Since speculation is linked to subjectivity. The result was that they got 92.37% for inference during clue detection in terms of f1 and 89.64% for negation. They got 84.07% for negation and 78.88% for inference during scope detection tasks in f1. And for rejection the PCRS was 80.26%, and for skepticism 71.43%. The investigation concludes that lexicon data and syntactic features require automated recognition of the clues to be performed to detect the reach of a keyword(Cruz Diaz et al., 2015)[31].

**A Machine Learning Approach to Sentimental Analysis in Multilingual Web Texts**

Boiy & Moens(2009)[32] performed experiments to characterize the nostalgic customer comments in blog phrases and sentences. Some articles are in French, Dutch, and English. The principal objective was to categorize the examinations into "positive" and "neutral" sentimental classes. In this case, they used different approaches to address the problems of the mission. Methods such as knowledge-gathering, natural language processing, and machine learning-enabled them to achieve best outputs. Unigram features are expanded with a limited number of language characteristics. The results are classified as 83% for English, 70% for Dutch, and 68% for French web information. They made a few observations and found that the efficiency is improved by filtering neutral phrases before implementing the sentiment classification algorithm. This error was noticed even in cases of sentence language that is substantially different from the formal language. The main cause of the error was incomplete training examples. And they also have to label training for other difficult tasks. However, the use of active learning strategies made this task simple. By integrating several approaches, active learning increases the F estimation over randomly selected labeling cases. A major advantage of active learning is the minimization in picking the examples and the quest for particular sentiments. There are several duplicates and unduplicated instances in the knowledge on the internet. This can be overcome by sentiments.

**Sentiment analysis using machine learning for business intelligence**

Chaturvedi et al., (2017)[33]suggests the use of the classification of sentimental analysis as an efficient way of analyzing textual data from different online platforms. Feeling analysis is nothing but a data mining approach which tests machine learning techniques for the consumption of textual data. With an enormously vast range of users' opinions, reviews, feedback, and suggestions available on the web, discovering, analyzing, and consolidating their views is so essential to improved decision-making. Opinion mining offers an accurate and dynamic view of customers in real-time and may have a huge impact on business decision-making. They also recognized potential areas for further research from the Business Intelligence research pond.



Electronic copy available at: https://ssrn.com/abstract=3674982

**A Survey on opinion mining and sentiment analysis; task, approaches, and applications**

Ravi & Ravi(2015)[34] involves the evaluation of emotional analysis of work from 2002-2014. It discusses 6 tasks such as classifying subjectivity, classifying emotions, calculating value, and generating lexicons, collecting opinions and product aspects as well as many opinion mineral applications. For these tasks, problems can be overcome quickly, and an objective outcome can be achieved. This paper is primarily concerned with including a list of public data sets for study of feelings. Examinations of the techniques were given to make use of SVM, NN, lexicon-based approaches immensely, but also finding that some techniques not exploited, such as random forest, evolutionary computation, association rule mining, rules miner, random field conditioning theory (CRF). The pattern vagueness can be done with fuzzy reasoning. Domain information can be given by CRF. Big data are used for RBFNN and for online learning algorithms. Through globalization, ontology can be utilized. Social network also has a massive part to play. The use of these enhanced techniques will obtain personal views extremely efficiently.

**Exploring the impact of Social Media on hotel Service performance: A Sentimental Analysis Approach**

Duan et al., (2015)[35] addresses the importance of social network in the use of emotional research on hotel services. Social media have become a big source of opinion, feedback, reviews and so on for users. It has become a customer and business knowledge center. A study of sentiments in the field of hotel services is performed to concentrate on an opinion of the individual concerning food, quality of service. The text mining method conducts this analysis in the evaluations received. User reviews are divided into different categories, based on the amount of service offered by the SERVPERF model. A new automated text mining method is used to detect consumers' opinions efficiently and accurately using the data generated by the online user. Further work was carried out into the five common types of user feedback. Such measurements are then combined with a model of econometric to form user opinions. The findings also demonstrate that different types of user feedback have different consequences. The value of social media in the industry is demonstrated.

**Machine learning, Sentimental analysis, and tweets: An Examination of Alzheimer’s disease Stigma on twitter**

Social background knowledge of aging isobtained from data sets accessible to the public. The method used is semi-automated text coding, machine learning methods for collecting AD tweets, and AD (API) keywords. Content checks are carried out on Twitter for Alzheimer's disease (AD) and Twitter's dementia representation. Talks about the value of social media, its structure, and how it leads to certain outcomes of growth through the use of these technologies. They have utilized interdisciplinary approaches to analyze social media data with basic programming skills. The researchers used customized methods for analysis. This research has shown that 21.13 percent of AD-related tweets have been evaluating public stigma using AD-related keywords. This social media is also a forum for attitudes and this stigma will contribute to aging and health conditions. AD and another dementia-related negatively impact people with dementia. Sensitivity to the power of social media should also be increased. It ultimately contributes to social impact initiatives(Oscar et al., 2017)[36].

**On Multi-tier Sentiment Analysis Using Supervised Machine Learning**

A multi-level classification architecture, consisting of important modules like cleaning the data&data pre-processing, function chosen, and classification training with a multi-level prediction model is proposed. It carefully describes the architecture and its parts. In experiments that test the efficiency of the suggested multi-level architecture with an analysis of the sensitivities and opinions of 150,000 film viewers, there are four classifications. The findings have shown that the multi-tier model increases the predictability over the single-tier model dramatically when the personalized dictionary is used by more than 10%(Moh et al., 2015)[37].

**Real-Time Sentimental Analysis on E-commerce Application**

Jabbar et al., (2019)[38] used the SVM computer training technology to examine beauty goods and musical instruments sentimentally. It solved problems by applying two types of categorization, e.g. categorization of test levels and categorization of sentence rates. Sentimental research strengthens product feedback marketing campaigns. Therefore, user reviews are important details about the products themselves. The users will learn about the products by star score, emoji, comments from the other consumer. It is an approach based on the sentimental orientation (SO), a kind of text classification



Electronic copy available at: https://ssrn.com/abstract=3674982

system. The advantage for the customer is that he or she can share his or her perspective on goods. It helps marketers to address the issue and to manufacture products according to consumer expectations and specifications. The data utilized for this investigation was obtained from Amazon website reviews. The real-time sentiment analysis system was introduced here. This is used by consumers and service providers in the e-commerce framework as time-saving. Build-in features and supporting libraries are used for assessing simple sentences' polarities. They also mentioned that machine learning and profound learning are used to manage complex structural phrases.

**An Automated System of Sentiment Analysis from Bangla Text using Supervised Learning Techniques**

Tuhin et al., (2019)[39] addressed the important role of sentimental analysis in the Bangladeshi text. Bangladesh is the world's fourth-largest country. Around 250 million speakers are in attendance. To derive the feelings from the Bangla text, two techniques namely Classifiers of Naïve Bayes and modern approach techniques were used. While comparing these two approaches, they found that the topical approach was better at both stages, as it was 90 % effective. Sentimental analysis is often referred to as data mining for textual emotions. This analyzed six types of feelings, such as positive, sad, delicate, excited, angry and scared. Substantial studies may be carried out at different rates. In the present research paper emotions are determined by the sentence and the magnitude of the documentation. Customer reviews are very important for areas such as analysis of goods, monitoring of social media, market research and analysis, etc. Nevertheless, there were still drawbacks that the regulated technique was not the best way of collecting enormous data. Because it's not easy to understand the data and Bangla language. More missing data were collected, so noisy data are not extracted from the Naïve-bays. For performance improvement, time-consuming data is needed.

**Using text mining and sentiment analysis for online forums hotspot detection and forecast**

Li & Wu, (2010)[40] researched on the identification of the hotspot and predicted online forums. In sentiment analysis, text mining is a significant technique. The emotive polarity of text is evaluated by the algorithm. The text can be analyzed to determine control and emotional polarity. K-means clustering and SVM groups are applied to evaluate the cluster for the online sports fora. An unregulated clustering algorithm is used for grouping forums into different clusters where hotspot forums in the center of each cluster are presented. The next time window is expected. Empirical research offers a relationship between post text feeling and hotspot distribution. To businesses, information-seekers, marketing teams, financial departments, etc. hotspot prediction approach is very useful as hotspot identification allows us to see what problems the company is worried about particular goods and services. Moreover, they conclude that it is not exact to predict hotspot using past data. Therefore, a supervised learning algorithm such as SVM is built to address this problem. There is also toxic extraction. An online server has also been completed. In many ways this allows users to obvert the distribution of the hotspot forum and natural rising b-groupings, to forecast the distribution of the hotspot forum, to choose how to view hotspot results, etc.

**Classification of sentiment reviews using n-gram approach of ML**

Tripathy et al., (2016)[41] reviewed the movie using a range of supervised algorithms such as Naive Bayes, Maximum Entropy, the Stochastic Gradient Descent, and Support Vector Machinery. It is proven, however, that the use of unigram, bigram, trigram, and the combination of these, and the mixture of TF-IDF and Counts Vectorizers as the combination to convert texts to a numerical matrix have shown better results of precise classification. Nevertheless, it is also disadvantageous that the Twitter message cannot be checked in limited amounts or in situations where certain statements or symbols reflect the feeling and repetition of the last letter several times. Both these weaknesses can be used to boost the recognition of feelings for potential research.

**Evaluation of sentiment analysis on smart entertainment and devices**

Gamal et al., (2019)[42] reviewed data sets on film review, product review, and smart electronic devices of five years by using common machine teachers. Simulations analyzes contain six sub tasking tasks, including sentiment classification, sentiment lexicon generation (SLG), sentiment quantification (SQ), opinion extraction (OE), function-based summative (FBS), and opinion spam (OS). After a comparative analysis, their work says that Support Vector Machine (SVM) and Naive Bayes (NB) are the most utilized algorithms for the classification of feelings in film and smart electronic scans. It used extractor is N-gram and its findings indicate that 90.3 percent of high precision is achieved by using the artificial neural networks (ANN) and the unicorn as function extractor.



Electronic copy available at: https://ssrn.com/abstract=3674982

**Managing Marketing Decision-Making with sentiment analysis**

Kauffmann et al., (2019)[43] address that the tools for sentiment-analysis analyze text data and then partition it into positive and negative data and then find polarity in the reviews. The paper uses Natural Language Processing (NLP), text data mining, and clustering approachesto get a new insight on product evaluation by consumers, which allows for product rankings based on product characteristics. The selection of features is often rendered by consideration of n-gram characteristics and the positive and negative characteristics. Through these tests, they have shown that this technique is indeed legitimate, by collecting vast quantities of data from amazon online reviews.

**Panoptical view of sentiment analysis technique**

Different methods and techniques such as machine learning techniques, Lexicon based methods, hybrid methods, Rules-based techniques, and ontologic methods have been investigated and compared. They used separate levels for opinion mining such as penalty, document level, sentencing level, aspect level, and emotional level. (Kakulapati, 2017)[44], says that a method depends on machine learning and an approach based on Lexicon can be the promising path to sentiment analysis. They conclude in their paper that everyone contributes directly and indirectly to mining and sentimental analysis has spread and is highly challenging in businesses. Here they addressed many studies, challenges, and analyses based on sentimental analysis and many issues.

**Predicting supervised machine learning performance for SA using contextual analysis**

The entire study of unstructured data including microblogs and feedback is performed. Their research says that one of the famous models of sentiment analyzes is a supervised learning algorithm which appliesdatasets and class labels depends on mathematical learning for training datasets, but that this model does not guarantee good output for data in real-time because of the variety of new data. The error has been found between 2.75 and 3.94 and value between 2.30 and 3.51 for different models of Ml of the average absolute difference. Besides, this work shows that contextual analysis can recognize words which willtransform from one dataset to another and lead to less prediction in the ML models, which automatically cluster the words into positive and negative words. In the future, however, the work must be undertaken to increase the results of the forecast when analyzing an increasing database in real-time(Aziz et al., 2019)[45].

**Sentiment analysis algorithms: evaluation performance of the Arabic and English language**

Abo et al., (2018) [46] gather multi-language data from different web forums and social networking sites, including English, Modern Standard Arabic (MSA), and Dialect Arabic (DA). Then they evaluated the opinion on three datasets bypassing them in Rapid Miner Tools then using machine-learning algorithms namely decision tree and naive Bayes classification, measuring the accuracy of both ML algorithms and resulting in English and Modern Standard Arabic having good results over dialect Arabic. ML algorithms need to improve accuracy by more than 50.80% to achieve better results for DA. Various ML techniques such as Vector Machine Aid, KNN, and profound learning for English or Arabic language assessment are stated.

**Sentiment analysis of Indian movie review**

The study has done opinion mining on Indian movie review corpus by applying a machine learning classifier. To perform testing, Bayesian classifier was used. The functional selection mechanics and the use of this classification were used to train characteristics of the corpus with the use of five feature selection algorithms, such as the chi-square, info-gain, gain-ratio, one-r and relief attribute, and comparative studies on these algorithms. And the results of this study show the maximum number of characteristics, the Relief F selection method with better F-value and low FP rate has been found to be successful (Trivedi, 2017)[47].

**Sentiment Analysis of YouTube data**

Bansal et al., (2019) [48] gives views on YouTube data. They used different data mining methods to extract data from YouTube given by the users for analysis. For feeling analysis, a Vader algorithm may be used but does not deliver the right result. The classification algorithms of machine learning called decision tree and K nearest neighbors and support vector machine were applied in the construction of a model and comparative analysis was performed using this algorithm. It allows SVM to decide based on the consumer satisfaction feelings analysis SVM is then used for more emotional analysis.

**Sentiment analysis techniques of twitter data**

Alsaeedi & Khan (2019) [49] use different methods of machine learning and a lexicon-based technique to



Electronic copy available at: https://ssrn.com/abstract=3674982

investigate opinions on Twitter info. And they found that SVM and MNB methods of machine learning provided the highest accuracy for several features and SVM classifiers were viewed as standards. For human-marked documents, Lexicon based technology was needed when it was feasible and little necessary. The study of the hybrid Twitter sentiment analysis algorithms appears to obtain classification precision of 85%, as it mixed multiple classifications, sometimes different feature types so that it's better than the supervised learning machine if it was done with an n-gram and bigram model, with Naive Bayes, Maximum Entropy and SVM of the machine learning. And the review of active learning technologies to identify twitter feelings for decision-taking in the future.

**Sentiment Analysis using machine learning for tweets about URI attack**

Twitter has used the most prominent social network where everyonecan share their thoughts about such events, and this data collection was prepared by various techniques that help to uncover feelings and feelings. This research was also used to analyze the Uri attack to find emotions and polarity about a specific incident. And finally, he concluded that all emotions for large quantities of tweets can be categorized by future big data(D. R. Kawade & Oza, 2017)[50].

**Sentiment analysis using speech data**

To find out how individual speakers are involved in a conversation to develops a model for evaluating feelings about speech data. Emotional analyzes is used to understand the mentality of people engaged in conversation and that to be able for a computer to understand human mentalities, they would know who is engaged in the talk and what the people who engage in the conversation are saying. A program was created that would know the identity of the speaker and evaluate the person's voice to detect motivation, as there could be more than anyone involved in the conversation, which would make it difficult to evaluate. Nevertheless, the downside of the model is that only one person can speak at a time to define a person's emotion; it cannot understand when two individuals speak simultaneously. And they conclude that they will develop the precise and scalable method in the future(M. S & Muthu, 2017)[51].

**Sentimental analysis of amazon product using ensemble algorithm**

The authors suggested a method for improving the precision of the analysis classification in this paper. Opinion mining is also a method of data investigation, which involves the compilation, interpretation, processing and evaluation of the review to the customer. They obtained data from various government websites in their paperwork. The collected data is pre-processed to delete unsolicited data, and trained data are sorted by using the Naive Bayes and SVM algorithm. However, these two algorithms are less accurate. For the proposed method, the combination of Naive bays, supporting vector machines, and ensemble algorithms are used, thereby ensuring the precision and speed of execution of the algorithm. So, after precision measurement, if the precision is high, the method will be used as a recommendation for the user. Then the standard evaluation so features for the algorithms are specified. The defined data set of certain items, such as electronics, books are readily read and categorized to provide good accuracy and performance, for instance in product review data set in Amazon. To suggest an item to the customer depends on customer feedback, the rating classification is widely used(Sadhasivam & Babu, 2019)[52].

**Machine Learning-Based Sentiment Analysis for Twitter Account Election Tweets**

In elections, various techniques of machine-learning and methods for opinion mining must be applied. The contribution is to tackle these issues admits the introduction of a blended methods which implies an investigator of emotions, like machine learning. This article compares sentiment analysis approaches for political analysis by using supervised algorithms such as Naïve Bayes and vector support machines (SVM). The study of political opinions is also discussed and (Hasan et al., 2018) [53] focuses on the adoption with machine learning algorithms of various sentiment analyzers to establish a highly accurate approach to learning about election feelings. The semitone orientation is the title, sentence, or phrase measured in a document for lexicon-based opinion mining. The machine learning techniques is mainly aimed at classifying algorithms such as Navies Bayes and SVM into the text format.

**A Machine Learning Approach to Sentiment Analysis in Multilingual Web Texts**

Boiy & Moens, (2009b)[54] presents our experiments on sensitivity analysis on the World Wide Web (WWW) for blogs, reviews, and online forum texts. The automatic feeling analysis includes extracting a feeling from an unstructured resource, including text, photos, or audio. For automated sentiment analysis, there are several further advantages. For example, public who expressed their opinions typically have stronger than average opinions. This paper introduces our expertise in the study of feelings expressed on the WWW in journals, comments and forums.It indicates an interest in people's thoughts about certain



Electronic copy available at: https://ssrn.com/abstract=3674982

consumer goods. Opinion identification is considered a supervised classification function. The purpose of the classification concerns web phrases such as blogs, newsgroups, forums, and review texts referencing a particular individual of interest. To gain insights into the classification of feelings in journals, customer feedback, and the newspaper forums, the authors performed a variety of experiments.The challenge was to correctly identify the terms according to positively, negatively, and neutrally expressed sentiment groups about an object of interest in a situation in which a relatively small amount of examples elucidated for training are given and the text is mostly unformed. The integrated approach incorporating knowledge processing approaches, natural language handling, and machine learning has produced best outputs because the job is difficult.

**Sentiment analysis for a product review.**

Sultana et al., (2019)[55] presents the theoretical study of several well-known sentiment analysis approaches. Certain approaches concentrate on verbs, nouns, adverbs and adjectives. Latest studies have proven that opinion mining is always best than the adjectives alone than the combination of adverbs, and adjectives. But the whole blend of verbs, adverbs, adjectives has not been focused. Six well-known supervised classifications are regarded as a newly introduced method, such as the speech pieces. This is also a mixture of adverbs, adjectives, and verbs that best incorporates speech parts.

**The Usage of Sentiment Analysis for Hearing the Voice of the Customer and Improving Businesses**

The analysis of sentiment and classification of user opinions is a collection of techniques, typically applied in computer software, which interpreted, measured, recorded, and used feelings, opinions, and emotions, which are generally referred to as impressions and contained in the web, social and business knowledge services. This research involves the analyses and classification of user opinions. The methods of sentiment analysis use technology to track views and to gain qualitative information, evaluate and report on what is in a positive. Microblogs are the most desirable data sources business parties choose to use. Microblogs like Facebook, where users share real-time feedback and thoughts on everything, face new challenges. Opinion mining is used to interpret the emotions of sentences, phrases, or documents through various methods. The Naïve Bayes, the full enterprise, and the help vector machine offer Natural Language Processing (NLP)(Axhiu et al., 2014)[56].

**Survey on sentiment analysis of the stock market**

The main aim of N. S et al., (2017)[57] is to discuss the different approaches used to interpret emotions. Social media are constantly used by people to exchange thoughts, search for feedback, etc. with the development of the new technologies. Therefore, social network views posted by people are better than those gathered from questioners who are usually full of uncertainty (reluctance) and lack personal interest. Different methods and algorithms for sentimental analysis may be used by the application and with the data. It uses Twitter, for example, to forecast stock prices for the future. A nostalgic stock market analysis helps people determine whether to invest in a company, with details. A review of stock refers to an overview of a company's market.

**Scientific Text Sentiment Analysis using Machine Learning Techniques**

Raza et al., (2019)[58] is mainly intended to allow researchers to assess the content. It uses an existing annotated corpus to examine the interpretations of science papers using quotation phrases. The noise was extracted from data using various data norms to clean up the data corpus. It uses six separate machine learning algorithms to perform the classification of this data collection. The following systems are being implemented using a Python-based machines learning libraries called Scikit Learn: Naïve-Bayes (NB), Support Vector Machine (SVM), Logistic Reversal (LR), Decision tree (DT), K -Nearest Neighbor (KNN) and Random Forest (RF). Scikit-Learn is a renowned Python language learning software library, offering an easy to use interface with m. First of all, the data stored in the file in format is read by our program. The output proved that the SVM and Naïve Bayes is performing best than the other classifier sets, according to a 60:40/60/40 ratio.

**Social Media Analysis for Product Safety using Text Mining and Sentiment Analysis**

(Isah et al., 2015)[59], developed a methodology to collect and analyze views and perceptions of drug consumers and beautifying items utilizing machine learning as well as text mining and opinion mining. Recent applications to the sentiment analysis about this research included crime surveillance, social network observing systems, five primary components that each observing system would cover, and a report on civilian responses, emotions, and communication in response to terrorist attacks was suggested as part of a framework. Opinion mining can be treated as the usual methods of the Lexicon-based



Electronic copy available at: https://ssrn.com/abstract=3674982

method, as explained in detail in the dictionary, using terms with their semantic orientations, are supervised semi-supervised and unattended classification tasks. The method based on learning is often represented in the training of a classifier to create a prototype. To show how machine learning can be used to infer feelings from social media data that display views and experiences of users.

**Web-Based Traffic Sentiment Analysis: Methods and Applications**

The methodology and implementations of web-based analysis of traffic feelings have been explored in this paper. The views of people have been taken from blogs, wikis, web forums, and social networking groups. Smart transport systems (ITS) were therefore unable to evaluate and proposed traffic sentiment (TSA) analysis for website traffic information. The study of the traffic sentiment was accompanied by different measures. For example: designing TSA's application architecture, building the interconnected bases for the TSA program, differentiating the merits and drawbacks of rules and learning-based techniques depends on the characteristics of website data suggested to use a rule-based algorithm for measuring sentimental polarity, taking into account changes in sentencing patterns They also say that the policy appraisal section will allow decision-making of managers. It helps assess public sentiment and address needs across all these strategies(Cao et al., 2014)[60].

**Feature selection for sentiment analysis based on content and syntax models**

(Duric & Song, 2012)[61],concentrates on feature selection for content-based and syntax-based sentiment analysis. The new opinion mining approaches are focused on feature selection methods in which human characteristics are collected. The characteristics are chosen using general statistical methods based on empirical evidence. The statistical method has one of the key advantages: it is fully automatic. The Model is trained to learn the features from the feedback automatically and to classify them into positive, negative and neutral comments. Subjective expressions are focused here, while entities are overlooked. With the advanced machine training methods, the outputs received by applying these functions in a highest entropy classification are competitive.

**Twitter sentiment classification using distant supervision**

Go et al., (2009)[62] shows that Twitter is one of the best microblogging services and has no previous classification for messages feeling. The algorithm for computer teaching classifies the feeling of Twitter messages with remote supervision. In the course of our training with emoticon data, the exactness rate is over 80%, compared to Naïve Bayes, SVM, and highest entropy in the machine learning algorithm. The famous micro-blogging service of Twitter is named. The status received by users tweets. Often these tweets share views on various subjects. Emoticons as noisy data labels are an efficient way to perform distant supervised research. The most reliable classification of the sentiment in Twitter messages is possible through automated algorithms. Twitter messages with specific features are used for tweeting with similar results by the master learning algorithm.

**A survey of data mining and machine-learning methods for cyber security intrusion detection**

Buczak & Guven(2015)[63] investigates the different types of cyber security intrusion detection miningthe data and machine learning approaches. The technique for machine learning or data mining where well known cyber data sets are utilized while data are an important source of knowledge for machine learning and data mining. The technology used to protect computer systems and networks is cyber security from the theft or loss of their hardware, software and electronic data and disruption and misdirection of their services. Firewall, antivirus software, and intrusion detection systems are used to track, recognize, locate, track, and report malicious behavior and policy violations. Cyber security. There are three different kinds of cyber analytics: misuse, anomaly, and hybrid. For machine learning and the data mining approach in cyber breach detection, testing, and training datasets are relevant. One of the key benefits of the ID method is that it can enter the data on the network and kernel level.

**Real-time sentiment analysis of twitter streaming data for stock prediction**

Das et al., (2018)[64] addressed the twitter streaming data research for market forecasts in real-time. In the present study an effort has been made to establish financial decisions, like stock prediction, to analyze or predict a company's potential stock prices. Twitter knowledge is worldwide one of the best micro-blogging services and is the most comprehensive source of online, public discourse. Streaming allows the real-time, programmatic, streaming data analysis of information is continuous streaming of information produced from various sources like websites, applications for mobile telephony, server logs, social media, trading floors, etc., which should be continuously processed through streaming without access to all the data. The key benefit is simplicity and flexibility that helps better evaluate and forecast user behavior in



Electronic copy available at: https://ssrn.com/abstract=3674982

an unabated way. Spark streamer is an extension of the Spark API core that permits scalable, high-performance, and fault-tolerant streaming of live data streams as Twitter. The classifying models are trained on historic data to deliver accurate results. Spark streaming provides abstraction at high rates.

**Detection and scoring of internet slangs for sentiment analysis using SentiWordNet**

Asghar(2014)[65] focuses on the identification and evaluation of Internet slags for SentiWordNet sentiment analysis. For sentiment analysis, information produced online is a significant source of data, including social network platforms, that extracts user opinions, sentiments, and feedback, views, and expectations about products and services that help business people learn about their customers and improve their business to satisfy their customers. Web slangs refer to the different slang forms used on the internet by various users, and due to frequent changes in its existence, it is difficult to describe the Internet. The result suggested that SentiWordNet can be utilized as a valuable tool for sentimental categorization jobs, where each word is correlated with numerical scoring that indicates positive and negative feeling knowledge.

**ELS - a word-level method for entity-level sentiment analysis**

Engonopoulos et al., (2011)[66]focuses on the ELS-an entity-level sentiment analysis approach used to model sequence with the use of CRF. To interpret the feeling in certain words of the list and to identify the feelings based on their person, they are qualified. In comparison with the CRF bag-of - word approach, Conditional Random Fields is considered to be the better approach for its sequential existence. Conditional Random Field is used to model the emotional tendency of webpages, which are divided into different types of comments like positive, negative, and objective commentary. CRF is used to identify feelings about a particular entity, blog post, or larger product review. A common way to create model sentiment analysis is to use a model bag-of-words to generate a document in vectors that allow each word in a document to score.

**Sentiment analysis applied to the educational sector**

Ravi et al., (2015)[67]discuss the application of opinion mining in the education field, which is frequently applied in different fields, such as political, industrial, medical, e-commerce and social media, etc. It includes collecting consumer feedback, thoughts, perceptions, and preferences of goods or services, which are a very valuable and important source of knowledge for corporate employees, to increase customer loyalty that results in income. The first step is data collection followed by text preprocessing, calculation, assessment, and eventually evaluation and visualization. The first and main step is data collection. Sentiment analysis can be used to obtain valuable information on a teacher's learning methodology and also on the curriculum. Analysis of feelings identifies the learning curve of students. The analysis understands the students ' needs. The analysis previews its performance.

**Detecting risks in the banking system by sentiment analysis**

Nopp & Hanbury(2015)[68]focuses on the identification of bank risks by analyzing feelings by the evaluation of Bank actions and risk opinions through the analysis of text data. Several studies have been performed to classify risks by text mining. The addition of sentiment indicators has shown a stronger predictor of financial distress. It includes information like CEO's letters and the annual report of banks and involves two distinct methods that are strong and highly important in the correlation between uncertainty and negativity within text-data, both as opportunities and as limitations for identifying risks in the banking system using sentiment analyzes.

**Sentiment Analysis of Social Networking Sites (SNS) Data using Machine Learning Approach for the Measurement of Depression**

Hassan et al., (2017)[69] presents how to discover a person's despair level by examining and deriving feelings from texts by applying theories of emotion, learning machines, and methods for the processing of natural languages on various social media platforms. It compared the sentence level feeling investigation for despair estimation between SVM, NB, and ME classifiers. It adopted the voting pattern and the chosen technology and scrutinize the results on two different datasets, twitter and 20 newsgroups. The examination shows SVM superior results by the classifications of Nave Bayes or Maximum Entropy. This is found that the SVM precision is 91%, the precision of the foundation of the navy is 83% and the precision of the maximum entropy is 80%.

**Sentiment analysis in Facebook and its application to e-learning**

Ortigosa et al., (2014)[70]presents a new SentBuk way to evaluate the sentiment on Facebook based on user-written posts. SentBuk collects and identifies the chats created by users of Facebook, helps the



Electronic copy available at: https://ssrn.com/abstract=3674982

identification of emotional change, identifying friends' emotions, classification of users by posts, and statistics among others. A hybrid approach follows the classification procedure implemented in SentBuk. The results from this approach show that sentiment analysis can be conducted in high precision on Facebook (83.27%). E-learning can be used to support personalized learning through adaptable e-learning systems, considering the emotional condition of the user in recommending the most appropriate activities for him or her. On the other hand, feelings towards a course may be used to provide teachers with input, especially about online learning, in which interaction with the person is less common.

**A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts**

A new machine-learning approach is proposed to assess this feeling polarity, using techniques of text classification only in the subjective portion. The extraction of these sections will be carried out by utilizing the effective methods for identifying minimal cuts in graphs and makes it easier to incorporate contextual constraints in cross-sentence. It examined the relationship between the observation of subjectivity and the categorization of polarity that shows that the detection of subjectivity can comprise assessments is far smaller extracts which still hold polarity data at a quantity comparable to a total evaluation. For the polarity classification of the Naive Bayes, the excerpts from subjectivity are more effective than the original document which states that they reflect not only the intended polarity but "clean" as well. Using the minimum cut system, efficient feeling analysis algorithms can be created. Through using this method contextual knowledge can increase the accuracy of polarity classification statistically significantly(Pang & Lee, 2004)[71]

**Network Environment and Financial Risk Using Machine Learning and Sentiment Analysis**

Li et al., (2009)[72]the GARCH-based ANN and SVM examines links between the volatility of financial commercial quantity and quantity of online data and therefore successfully forecast the network financial risk. All approaches can yield good results for prediction. The GARCH-based ANN predicts the volatility trend rather than the GARCH-based SVM, while the GARCH-based SVM predicts the volatility itself with the GARCH-based ANN. There is also a strong link between the sense of information and volume of trading as well as volatility in the asset price. The analytical studies are recommended and funded to provide an alternate way to forecast volatility and trend portfolio holders, academics, etc.

**Web Application for Sentiment Analysis Using Supervised Machine Learning**

Laryea et al., (2015)[73] implemented thus a five-pronged system of positive, negative, suggestion, no emotions, and sentence neutral. They shattered data from amazon.com and used software for the processing of open-source language to get their feeling out. They also showed how the places to review the feeling map and the data for our classifier’s training. The excessive number of classrooms given improved the performance of the classifiers for positive, negative, and non-sensitive classes and very low for neutral and advisory classes.

**Sentiment Analysis of Textual Reviews Evaluating Machine Learning, Unsupervised and SentiWordNet Approaches**

V. K. Singh et al., (2013)[74] presents experimental findings in the assessment of the output of all three methods for classification of feelings at the document level. It has implemented two classifiers based on master learning (Naïve Bayes and SVM), the SO-PMI-IR algorithm, and SentiWordNet approaches to classify the sensitivity of the script. Two preexisting large datasets were used and one of moderate size was collected by ourselves. The paper makes two valuable contributions: (a) it provides a detailed performance evaluation of all three existing film review methods, and (b) it introduces a new updated SentiWordNet approach Adjective and Adverb combination scheme.

**A domain transferable lexicon set for Twitter sentiment analysis using a supervised machine learning approach**

Ghiassi & Lee(2018)[75]suggested a new method towards the explanation for feature selection of a lexicon specific to Twitter Sentiment Analysis (TSA). For this analysis, the character set is reduced to seven "component characteristics" to decrease sparsity. Based on these characteristics, the Twitter Specific Lexicon Collection (TSLS) can yield very precise results. It uses a dynamic neural network architecture (DAN2) and SVM as calculated by the reminder, accuracy and F1 metrics Theoutcome proved that the Twitter Generic Feature Set (TGFS) extracts from 2 different datasets (@JustinBiebers and @Starbucks) is transmittable in the domain and can generate excellent sentiment classification values combined with just a Twitter Domain Specific Features (TDSF).TGFS is evaluated for its efficiency and



Electronic copy available at: https://ssrn.com/abstract=3674982

transferability over three different areas (@GovChristie, @SouthwestAire, and @VerizonWireless).

**Performance analysis of supervised machine learning techniques for sentiment analysis**

Samal et al., (2017)[76] uses seven promising supervised machine learning algorithms to build a simple but new approach to the analysis of feelings of the film reviews. They collected data sets from the different sizes and selected some of the common machine-learning algorithms, which are commonly utilized and supervised, to train the pattern or model. This will allow the model to identify the analysis. For processing film analysis, the Python NLTK software along with Win Python and Spyder are used. The sklearn module of Python is then used to exercise the model and to detect its exactness. For a great number of films, review results conclude that linear SVC / SVM is among the best classifier to achieve 100% precision.

**Price Movement Prediction of Crypto currencies Using Sentiment Analysis and Machine Learning**

Valencia et al., (2019)[77] the performance of three prediction models was evaluated and contrasted, using Facebook, market data, and RF data for Bitcoin, Ethereal, Ripple, and Lite Coin. Results show that crypto currency markets can be forecast with machine learning and sentiment analysis to predict those crypto currencies using Twitter data alone and that neuronal networks over-perform other models.

**Sentiment Analysis of Malayalam film review using machine learning techniques**

Nair et al., (2015)[78] demonstrates how sensational analysis can be applied using machine learning methods to analyze Malayalam films. This method is mixed, consisting of techniques for machine learning and a method focused on rules. This research will help the users evaluate the criticism of the film and give new films the rank and popularity. Two statistical methods for evaluating Malayalam film reviews were compared, namely SVM and CRF. The device controls the polarity at the sentence level in these works, resulting in a 91% precision.

**Sentiment Analysis on Movie Reviews using Information Gain and K-Nearest Neighbor**

In the analysis KNN with information gain features to achieve good performance will be evaluated using polarity v 2.0 from the Cornell film review data collection. Such studies were intended to find the optimum K for KNN and to use other approaches to compare KNN. KNN is the highest output approach with 96,8% accuracy compared to NB, SVM and Random Forest using the information gain function selection and the optimum K is 3 (Daeli & Adiwijaya, 2020)[79].

**Sentiment Analysis: Towards a Tool for Analyzing Real-Time Students Feedback**

Altrabsheh et al., (2014) [80] suggested an automatic analysis of feedback with opinion mining. The investigation of feelings is domain-based and has not been used earlierto provide real-time feedback, even though it was used in the educational sector. To find out the betterpattern for automated investigation, 4 aspects are examined: pre-processing, functionality, machine learning approaches, and the common class use. The highest outcome for all four was the SVM with the highest pre-processing, unigram, and no common class rates, which received 95 percent precision.

**Scope of negation detection in sentiment analysis**

Dadvar et al., (2011)[81] study the issue of the polarity of feelings in movie reviews when negative terms such as the sentences do not occur or can rarely happen. They used term frequencies to test the system's discriminatory ability in various window sizes. Classification accuracy can be improved by sarcasm and metaphor identification. The findings show that although different window sizes have been implemented, no substantial difference is made in classification accuracy. Nevertheless, the detection of negation led to the discovery of further expressions of opinion or sentiment. They conclude that traditional methods for negative detection are insufficient to analyze sentiment in this field and that progress needs to be made by exploiting information on the implicit expression of opinions.

**Exploring Impact of Age and Gender on Sentiment Analysis Using Machine Learning**

S. Kumar et al., (2020) [82]analyzes the sentimental impacts will encourage e-commerce companies, on the basis of different demographics, to market their items. The datasets are generated through the aggregation of Facebook user reviews and by asking them to answer questions about their interests in accordance with their age group and gender information. The data for each age group and sex was analyzed instead. In the end, feeling research uses many methods of Machine Learning (ML) to analyze the influence of age or sex on user input, including maximum entropy, a supportive vector machine, a convolutionary neural network and a long-term storage. In terms of gender, the Above 'Age 50' was the most reliable recorded by women for ages relative to all other age classes.



Electronic copy available at: https://ssrn.com/abstract=3674982

**Local Decision Pitfalls in Interactive Machine Learning: An Investigation into Feature Selection in Sentiment Analysis**

The effect of local interactions in IML is discussed in the sense of small-level text categorization features. It analyzed the decisions taken by users and their effect on model output with a mix of people-subject examination and reproduction. Through sentiment analysis, although promoted and recommended in IML research, it is noted that local interactive app selection is not as beneficial as one would wish. It discusses how the dataset, the learning algorithm, and the practice scale impact these patterns. They have clear generalization issues with these variables. The findings suggest that quick iterating with IML systems can be risky if the local steps are allowed to degrade the overall output and are isolated from the global background. The article ends by discussing the implications of our findings for IML systems and works in a wider context(Wu et al., 2019)[83].

**Evaluating Machine Learning and Unsupervised Semantic Orientation approaches for sentiment analysis of textual reviews**

Waila et al., (2012)[84] introduces an experimental estimation of machine learning categorization approaches using the SO PMI-IR algorithm for opinion mining of film review documents based on Unattended Semantic Orientation. It used our dataset array and preexisting data sets, containing a huge volume of Hindi film customer reviews. The approaches to Naïve Bayes and SVM were carried out in several ways. Besides providing a comprehensive comparative view of these techniques, the results show that the Naive Bayes algorithm has a relatively good range of features and often corresponds to the common SVM high efficiency, for sentiment analytical research at least. Without the need for a previous preparation, the SO-PMI-IR algorithm produces significantly exact feeling classifications. However, the precision of SO-PMI-IR lean on POS tags that are utilized as characteristic and thresholds.

**A personalized recommender system using Machine Learning based Sentiment Analysis over social data**

The proposed social network collects user feedback, restaurants' opinions and points of interest including activities, and places of interest, to configure and categorize recommendations based on user interests. Because the data is not sufficient for a fine-grained auto-conscious system. For further optimization of search query results, machine learning and opinion mining based methods are utilized. It gives users faster and more appropriate information so that non-relevant data is avoided and personalization is much needed (Ashok et al., 2016)[85].

**An ensemble approach to stabilize the features for multi-domain sentiment analysis using supervised machine learning**

Ghosh & Sanyal (2018)[86]explores the impossibility of individually using the commonly used system for selecting characteristics and their mixed solution to 4 classification algorithms for machines. Three basic data sets are used to test the approaches proposed. Dataset of IMDb film review, technology, and food analysis. Select function subsets from three separate selection methods in the first case. Such a method is then used to combine these various function subsets using statistics UNION, INTERSECTION, and a revamped UNION system to achieve all the top rankings including commonly selected features. Finally, use this function vector for classification of the revision data set to train classificatory SMO, MNB, RF, and LR. Estimation techniques such as accuracy, retrieval, F-calculate, and ROC curve calculate the efficiency of the algorithm. The integrated approach has been shown by experimental tests to be the best alternative 92.31 with the SMO classification that motivate and comparable to the related work.

**Fuzzy rule- based systems for interpretable sentiment analysis**

H. Liu & Cocea(2017) [87]suggests the use of furious rules as conceptual models to examine emotions precisely and interpretably. The practice of fuzzy logic is more closely associated with intrinsic language ambiguity, while the "white box" aspect of rules-based learning methods increases the interpretability of the tests. The approach suggested is evaluated on 4 datasets of film reviews to contrast its success in accuracy with two other sentiment analysis approaches that are well known for their results. The findings show that the fuzzy rule-based methods achieve considerably best than the common techniques of machine learning, minimizing computation difficulties, and maximizing interpretability.

**What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach**

Hew et al., (2020)[88] used the satisfaction of learners to assess performance with MOOC and expands



Electronic copy available at: https://ssrn.com/abstract=3674982

the theoretical insight into the factors influencing MOOC fulfilment. It includes three subscription to the field: (a) quantitatively evaluated data on 249 inconstantly chosen MOOC, and (b) a single technical system to investigate a broad dataset of consumer-produced feedback from 6 391 users. (c) Define specific variables of students and courses that can predict and estimate the relative effects of MOOC student satisfaction. The strength of the findings was reinforced by the scale of the analyzed data set, and the use of sampling probability (random sampling procedure) that can help generalize the results. This research explores particular variables at the learner and the course level that will forecast the fulfillment and related effects of MOOCs.

**A sentiment analysis model for hotel reviews based on supervised learning**

Shi & Li, (2011)[89] focuses on real-time restaurant reviews and proposes a supervised machine-learning method with 2 kinds of information to identify documents in polarity. The test results show that TF-IDF knowledge is more effective than frequency.

**Sentiment analysis of Facebook statuses using Naive Bayes classifier for language learning**

The central and underlying premise is that it can be viewed as a classification activity to understand how the human think about such subjects. This explores a new type of data display and notes major changes in the unigram patterns. This information is a random sampling of Facebook status streaming and has not been obtained with precise questions. The measurements of the hand-labeled data helps us to conduct cross-validation testing and detect differences in classification output between sockets. The experimental results show that precision with the Naive Bayes Classifier to evaluate the emotional state of Facebook users is very high(Troussas et al., 2013)[90].

**News-based sentiment analysis in real estate: a machine learning approach**

Hausler et al., (2018)[91] explores the connections between news-based feelings captured by an approach to machine learning and the securitized and direct US markets for commercial Immobility. It found the built sentiment indicators using a vector autoregressive frame to predict the complete returns of both markets. Even after we have reviewed macroeconomic factors and other proven feeling agents, the results show a leading relationship to the situation. In fact, empirical evidence shows that the indirect market has a shorter response time than the direct market.

**Sentiment Analysis of Review Datasets Using Naive Bayes and K-NN Classifier**

Dey et al., (2016)[92] focuses on sentimentally oriented Internet crawling platform for the fast invention and investigation of sentimental material of film reviews and hotel reviews. This used statistical methods to capture subjective style elements and the polarity of a sentence. Here, the supervised machine learning algorithms K-Nearest Neighbor (K-NN) and Naive Bayes are explored within great detail and their accuracy, precision, and recall values are compared. In film reviews, Naive Bayes generated 80 percent of the results of K-NN, but these algorithms were less accurate, almost similar in hotel reviews.

**Aspect based sentiment analysis of student's opinion using machine learning techniques**

Sivakumar & Reddy(2017)[93] suggested a new approach to analyzing input from Twitter API students through the calculation of semantic relations between word aspect and phrase of student opinion. Sentiment analysis with k-mean and naïve Bayes algorithms was performed. For the polarity of the sentences and the measured semantical relation between the sentiment sentence and a given aspect word, they used sentiment kit R. The findings were used to identify the individual phrase with a word dimension. The accuracy, recall, and F-score metrics achieved were good results. The findings may assist the users to develop their learnings and allow the teachers to enhance their teaching qualifications.

**A Sentiment Analysis Model to Analyze Students Reviews of Teacher Performance Using Support Vector Machines**

Esparza et al., (2018)[94] present a social analysis model utilizing a corpus of real Spanish comments on the assessment of tutors results. The software uses three kernels: linear, radial, and polynomial support system algorithms to find a positive, negative, or neutral classification of comments. As assessment methods, they measured sensitivity, specificity, and forecast values. The findings of this work will assist other studies enhance the interpretation of comments and recommend courses for tutors to improve.

**Sentiment Classifier and Analysis for Epidemic Prediction**

Das Adhikari et al., (2018)[95] offers a noble method for predicting the region susceptible to disease by using text analysis and computer education. A significant number of preprocessing of unstructured data is performed to turn it into a structured one, as well as various extraction techniques, such as count vectorization, TF-IDF, subject modeling, and so on. It uses the combination of words and grams, word



Electronic copy available at: https://ssrn.com/abstract=3674982

embedding, and TFIDF with various data extraction and profound learning algorithms such as SVM, Naive Bayes, and RNNLS. In contrast with others, Naïve Bayes has done well with TF-IDF.

**Effective Sentiment Analysis of Social Media Datasets using Naive Bayesian Classification**

Gurkhe et al., (2014) [96]attempt to determine the polarity of social media datasets with the queried keyword (positive, negative, or neutral). The following method provides an methods for spontaneously categorizing the feelings of social network information: first of all the practicing data are insert to the opinion mining engine to be trained by the machine learning algorithm. Once the learning has been done with trained precision, the system begins receiving separate social data to the keyword that it analyzes and interprets.

**Domain-based sentiment analysis in regional Language-Kannada using a machine learning algorithm**

Rohini et al., (2016)[97] aims to examine domain sentiments in regional film-specific language using a classification machine learning algorithm and to compare research using direct Kannada dataset with English translated machines. Opinion mining in Kannada is an incredible activity and only a few resources such as Training data for domain-specific, part of the Tagger speaker is available. Some words on machine translation in Kannada created an ambiguous text that is one of the key causes of inappropriate performance. In contrast to Computer Traduced English, the study of test data in the regional language performs better.

**Sentiment Analysis for Financial News Headlines using Machine Learning Algorithm**

Shuhidan et al., (2018)[98] deals with the implementation of an algorithm for the investigation of financial news headlines in Malaysia. This finding can be utilized by stakeholders who like to get update in financial news and check for financial information or data. The information is gathered from the online financial news of Malaysia, which comes from the New Straits Times business section. As a method to perform emotional research, it uses the Opinion Lexicon's algorithm and Naive Bayes' algorithm. This work includes various pre-processors such as data removal, word removal and clean-up of the dataset before the sentiment analysis is performed using the selected machine-learning algorithms. tm in R is used to clean the data set while extracting words, while the snowball stemmer is used for root word settings of the data.

**Microblogging sentiment analysis with lexical based and machine learning approaches**

Maharani(2013)[99] centered on evaluating the process by machine-learning using lexical and model-based approaches to identify tweets containing opinions using both methods. The tested approach is Vector Machine Support (SVM), Maximum Entropy (ME), Multinomial Naive Bayes (MNB) and Nearest Neighbor (k-NN) (SVM). The outcome of the test depends heavily on a lexical approach that was translated into a classification matrix for Opinion. Due to its capability to create new training data models based on the results model, a machine learning approach will deliver better accuracy.

## 3.10 USE CASES OF SENTIMENTAL ANALYSIS

Opinion mining is applied in different industries. Although the fields of application for sentiment analysis are related, they all enhance efficiency by analyzing public opinion changes ("Sentiment Analysis," 2018.)[100].

**Brand monitoring:** Sentimental brand analysis helps you to stay abreast of your market credibility, detect current or potential reputational problems, and respond quickly.

**Competitive research:** Able to observe and examine how society assesses rivals, just as you assess their role to any company. How do other market leaders expect most from customers? Was something lacking or wrong with competitors? What platforms do consumers use to communicate with other firms? Use this information to develop the contact, marketing campaigns, services overall, and to provide customers with services and products.

**Flame detection and customer service prioritization:** To maximize the work of the consumer services, hospitality brands, financial institutions, supermarkets, transport, and other companies are using sentimental grouping. The classification of incoming customer service messages by polarity, style, dimension, and priority can also be automated by text analysis users. As the spark should be put before it becomes fire, new messages are being processed first from the least satisfied and most angry clients.

**Product analysis:** Effective firms develop a minimally viable product (MVP), obtain early feedback, and continually enhance a product even after its publication. Data are given by surveys, social media and



Electronic copy available at: https://ssrn.com/abstract=3674982

forums, and customer service experiences. Issues such as identifying the consumer groups to query, reviewing, and classifying comments occur in this data ocean.

**Market research and insights into industry trends:** Analysis of emotions addresses the retrieval issue of vast quantities of unstructured data. Through using these text analyses, advertisers monitor and research patterns of customer behavior in real-time to forecast potential behaviors and assist management in making informed decisions.

**Workforce analytics/employee engagement monitoring:** Specialists automate the study of SA tech employee surveys, helping them to resolve issues and complaints more quickly. The total ton of comments, group outcomes by departments, and keywords can be identified and monitored by personnel managers, and how employee emotions have changed over time. Sentiment analysis gives the next phase of workers mood management with real-time tracking capabilities.

## 4. SUMMARY OF RELATED WORK :

**Table 1**: Review of findings presented by different authors between 2009 and 2020.

| SNo | Author(s) | Year | Inventions/Findings/Results |
|---|---|---|---|
| 1 | Boiy, E., & Moens, M.-F. [25] | 2009 | Classifying clients' emotions by means of approaches such as information processing, natural language therapy and machine learning systems, in English, French and Dutch helped them produce successful results. |
| 2 | Boiy, E., & Moens, M. [54] | 2009 | Integrated approach incorporating knowledge, information retrieval,recovery approaches, natural language processing and machine learning have shown strong results |
| 3 | Go, A., Bhayani, R., & Huang, L. [62] | 2009 | Describes the pre-processing steps to achieve high precision or accuracy. |
| 4 | Yang, C., Tang, X., Wong, Y. C., & Wei, C.-P. [35] | 2010 | This sentimental analysis compares and describes the favorite goods that make customers happy and looks at unsupervised approaches to learning for sensational study. |
| 5 | Li, N., & Wu, D. D. [36] | 2010 | Empirical research relates the feeling of post text to the distribution of hotspots. SVM is used for hotspot prediction with exact data tests. |
| 6 | Kumar, A., & Sebastian, T. M. [9] | 2012 | Examines the significant study of emotions and discusses their fundamental terms, tasks and granularities, functional and future applications and challenges |
| 7 | Duric, A., & Song, F. [61] | 2012 | Learn the features of feedback automatically and identify them into positive, negative and neutral comments. |
| 8 | Vohra, S., & Teraiya[2] | 2013 | Exploring and explaining its policies in this field by contrasting the concept of emotional research in natural language processing. |
| 9 | Medhat, W., Hassan, A., & Korashy, H. [5] | 2014 | Gives the overview and review of the latest SA algorithms and software updates. |
| 10 | Axhiu, M., Veljanoska, F., Ciglovska, B., & Husejni, M. [56] | 2014 | Stimulus studies are joined together for tracking views and collecting qualitative information, for calculating and communicating in a palatable way. |



Electronic copy available at: https://ssrn.com/abstract=3674982

| | | | |
|---|---|---|---|
| 11 | Cao, J., Ke, Z., Wang, H., Cheng, J., Qiao, F., Wen, D., & Gao, Y. [60] | 2014 | Overview of applications of TSA, creation of similar bases for the TSA program, contrasting both rules-based and learning based approaches' advantages and disadvantages based on web data characteristics |
| 12 | Cruz Diaz, N., Taboada, M., & Mitkov, R. [24] | 2015 | A lexical knowledge and syntactic functions is adequate to identify the reach of a keyword immediately recognizes the clues. |
| 13 | Ravi, K., & Ravi, V. [27] | 2015 | Explores the significance and vagueness of the model using fluid logic of SVM, NN, and lexicon based approaches. |
| 14 | Duan, W., Yu, Y., Cao, Q., & Levy, S. [28] | 2015 | The effective and exact detection of customer opinions on the data generated by the online user is conducted through a modern automated texture mining method. |
| 15 | Patil, H., & Atique, M. [34] | 2015 | A wide variety of methods for opinion analysis such as user opinion estimation, change in polarity. The classification of feelings, qualitative and broad-scale data analysis, cross-domain sentiment rating, perception of feeling variance, sensitivity, the identification of feelings, sales forecast, etc. are discussed here. |
| 16 | Isah, H., Neagu, D., & Trundle, P. [59] | 2015 | Machine learning approaches are applied to deduce feelings from social network data indicating drug users' thoughts and experiences |
| 17 | Buczak, A., & Guven, E. [63] | 2015 | Different types of intrusion detection, data mining and machine learning methods are summed up |
| 18 | Aydogan, E., & Akcayol, M. A. [6] | 2016 | Machine-based sentiment analysis is analyzed, identified by their knowledge extraction tasks and the possible problems are addressed. |
| 19 | Jain, A. P., & Dandannavar, P. [4] | 2016 | In the text analysis system for twitter data using Apache sparks, Naïve and Decision Trees Master Learning Algorithms are utilized for sentimental analysis |
| 20 | Tripathy, A., Agrawal, A., & Rath, S. [38] | 2016 | Study films with multiple supervised algorithms of machine learning, which are then used using an n-gram method. |
| 21 | Ahmad, M., Aftab, S., Muhammad, S. S., & Ahmad, S. [7] | 2017 | Machine Learning methods and methods for emotion analysis and classification are used for many experiments and studies. |
| 22 | Swathi, R., & Seshadri, R. [10] | 2017 | Understand recent work on Big Data Machine Learning, devices and systems used, and proposed methods and spaces. |
| 23 | Oscar, N., Fox, P. A., Croucher, R., Wernick, R., Keune, J., & Hooker, K. [29] | 2017 | Interdisciplinary methods for the study of social media data using simple programming (ML in Python) have shown that 21.13% of AD-related tweets use AD-connected keywords to evaluate public stigma. |



Electronic copy available at: https://ssrn.com/abstract=3674982

| | | | |
|---|---|---|---|
| 24 | Kawade, D., & Oza, K.[33] | 2017 | Addressed the value of TL and NPL for sentiment analysis, and concluded that the interest in English is less than in other languages due to lack of resources and research. |
| 25 | Gopu, M., & Swarnalatha, P. [37] | 2017 | Sentiments are specifically categorized by machine learning algorithms such as bag-of - words, n grams, naive grades, and natural language processing. |
| 26 | Kakulapati, V. [41] | 2017 | The topic of research, problems and analysis are explored in the sense of sentiment analysis. It also answered questions about sentiment analysis |
| 27 | Trivedi, S. [44] | 2017 | The maximum feature number is determined to be strong with best F-value and lower FP levels for relief - F feature selection approach |
| 28 | Kawade, D., & Oza, K. [47] | 2017 | An application for a study of Uri's attack was used to identify the emotions and polarity around particular events. |
| 29 | S, M., & M, R. [48] | 2017 | Develop a model which analyzes feelings on speech information to identify the emotions in a conversation of individual speakers. |
| 30 | Naiknaware, B., Kushwaha, B., & Kawathekar, S. [51] | 2017 | Using machine learning technology, a vast number of data can be analyzed to ensure that decisions are optimized and strategic. |
| 31 | Singh, J., Singh, G., & Singh, R. [53] | 2017 | Used three manually annotated datasets in four classifiers for sentiment analysis. |
| 32 | S, N., M, A., & K, A.[57] | 2017 | Use Facebook to forecast future share prices and bonds' emotional analysis, social media platforms will help people make intelligent decisions |
| 33 | Mäntylä, M. V., Graziotin, D., & Kuutila, M. [8] | 2018 | The automatic text clustering with manual qualitative analysis is possible to conduct a computerized literature review. For 6,996 papers, a bibliometric study of sentiment can be performed. |
| 34 | Mahendran, N., & Mekala, T. [12] | 2018 | A review of the available research methods, the issue and the application in the field of sentiment analysis |
| 35 | Abo, M., Shah, N., Balakrishnan, V., & Abdelaziz, A.[43] | 2018 | The research by moving the opinion mining into Rapid Miner Tools in 3 datasets and using algorithms for machine learning, measured the precision of both ML algorithms |
| 36 | Hasan, A., Moin, S., Karim, A., & Shamshirband[52] | 2018 | Focuses on the adoption of multiple Opinion mining with machine-learning algorithms in order to decide how to learn about election feelings at the most exact pace |
| 37 | Shirsat, V., Jagdale, R., Shende, K., Deshmukh, S. N., & Kawale, S.[17] | 2019 | Sentence-level study of news stories and blogs to measure the polarity of news and forum posts by Naive Bayes, the vector supporters and the random forest classification. |
| 38 | Bansal, A., Gupta, C. L., & Muralidhar, A.[26] | 2019 | Analyze YouTube feedback using various classification tools, including the decision tree, K-Nearest Neighbors and vector supports for strong predictions |



Electronic copy available at: https://ssrn.com/abstract=3674982

| | | | |
|---|---|---|---|
| 39 | Kauffmann, E., Peral, J., Gil, Ferrández, A., Sellers-Rubio, R., & Mora, H.[30] | 2019 | The UGC provides a great deal of knowledge to help marketing decision-making. Through means of nostalgic analyses, data mining and NLP techniques, they evaluate user feedback and increase sentence granularity. |
| 40 | Jabbar, J., Urooj, I., JunSheng, W., & Azeem, N. [31] | 2019 | Sentimental orientation (SO) of opinion is a text classification machine learning technique used for the management of complex structural phrases. |
| 41 | Tuhin, R. A., Paul, B., Nawrine, F., Akter, M., & Das, A.[32]. | 2019 | Emotions may depend on the volume of the expression and the nature of the text |
| 42 | Gamal, D., Alfonse, M., El-Horbarty, E.-S., & M.Salem, A.-B.[39] | 2019 | Artificial Neural Networks (ANN) and the unigram as an extractor achieve a high accuracy of 90.3%. |
| 43 | Kauffmann, E., Peral, J., Gil, Ferrández, A., Sellers-Rubio, R., & Mora, H. [40] | 2019 | Using NLP, text data mining and clustering techniques to gain new score on customer product feedback |
| 44 | Aziz, A., Starkey, A., & Madi, E.[42] | 2019 | Context analysis is used to classify certain terms shifting from one dataset to another and to make the ML models less predictive. |
| 45 | Bansal, A., Gupta, C. L., & Muralidhar, A. [45] | 2019 | Several data extraction techniques have been used to extract data from YouTube provided by the users for feedback. |
| 46 | Alsaeedi, A., & Khan, M. [46] | 2019 | Twitter data sentiment analysis is performed by various approaches and techniques of machine learning. |
| 47 | Liu, R., SHI, Y., JI, C., & JIA, M. [49] | 2019 | It summarizes the research outcome and focuses on algorithms and transfer learning applications in sentiment analysis and growth trend analysis. |
| 48 | Sadhasivam, J., & Babu, R. [50] | 2019 | The proposed method is used with a combination of naive algorithms, svm and ensemble, which discuss precision and speed of the output of the algorithm |
| 49 | Sultana, N., Kumar, P., Patra, M., Chandra, S., & Alam, S. [55] | 2019 | The theoretical analysis of some well-known sentiment analysis approaches, based on noun, adverbs, adjectives |
| 50 | Raza, H., Faizan, M., Hamza, A., Mushtaq, A., & Akhtar, N. [58] | 2019 | Sentiment analysis is performed by applying an existing constructed annotated corpus, using quotation sentences. |
| 51 | Yi, S., & Liu, X.[1] | 2020 | Apply Machine Learning algorithms for learning, analysis and classification on the basis of customer experience of product data and shop knowledge. |
| 52 | Yogi, T. N., & Paudel, N. [11] | 2020 | Comparative tests of classification algorithms based on machine learning are tested for sentiment analysis using various measurement parameters such as range, range, retrieval and F calculation in the three different data sets of different sizes |



Electronic copy available at: https://ssrn.com/abstract=3674982

| | | | |
|---|---|---|---|
| 53 | Suryawanshi, R., Rajput, A., Kokale, P., & Karve, S. S. [14] | 2020 | The methods of natural language processing help to evaluate the tweet emotions, where they are positive, neutral, negative and then graded according to the emotions. |
| 54 | K. Sentamilselvan, D. Aneri, A. C. Athithiya, P. Kani Kumar. [15] | 2020 | The classification and logistic regression of Naïve Bayes was used to evaluate feelings and classify according to the better precision of the scientific classification |
| 55 | Godara, N., & Kumar, S.[16] | 2020 | Twitter dataset output is counted by applying different efficiency metrics such as exactness and accuracy, retrieval and f estimation, including decision-tabing and artificial neural networks (ANN), naïve bays, fuzzy techniques (FU), and vector support (SVM). |
| 56 | Nihara Yadav, Ms. Sonal Arora. [18] | 2020 | Sentiment analysis can be used with supervised machine learning for fake positive or negative review identification. SVM algorithm not only passes text detection but can also be used for fake feedback |
| 57 | Gujar & Pardeshi. [19] | 2020 | Sets the path to social media opinion mining through the common machine learning approach by means of the Twitter API |
| 58 | Arote Rutuja S., Gaikwad Ruchika P., Late Samidha S., Prof. G. B. Gadekar. [20] | 2020 | Different aspects of the sentiment analysis were studied with a focus on Czech. |
| 59 | Raju, S. M., & Tarif, A. M. [21] | 2020 | Determining Bitcoin's predictable price path in USD by machine learning and emotional research. |
| 60 | Ardianto, R., Rivanie, T., Alkhalifi, Y., Nugraha, F. S., & Gata, W. [22] | 2020 | In evaluating e-sports sentiment for student learning curricula, calculating views, or distinguishing positive and negative feelings for e-sport education and Naïve Bayes algorithms can best be forecasted. |

## 5. DISCUSSION & FUTURE WORK :

A review of the applications of feelings analysis applicable to their methods and technologies are provided here. This analysis of study and literature offers valuable insights into the advantages and value of interpreting and classifying feelings. The primary issue is the exact classification of all sentiment analysis techniques. Many experiments have had high success rates, but the accuracy percentages in all approaches and languages have still not been achieved. Most sentiment analysis methods depend on English-language text, and the methods developed are therefore mainly English. Work should be improved in other languages because the knowledge is not accessible in English alone. The limited focus of this study is that the application side there is primarily focused on the exclusion of the hardware and theory-related aspects. The growth in deep learning is making emotional research more and more difficult. Sentiment analysis has just gone beyond a modern-technology curiosity and will shortly covert an important resource for all modern-day businesses. You particularly need to find the right Feature Set for current hot deep learning that may become the subject of the research on sentiment analysis.



Electronic copy available at: https://ssrn.com/abstract=3674982

## 6. RESEARCH GAP :

The traditional models of study are checked with data from the public review. This form of analysis, however, has not been closely examined with respect to feelings. This does not include original and secret knowledge outside real definitions of feelings by classifying the feelings as positive or negative. Other unique phrases are very complex and cannot be reliably categorized. Several weaknesses such as assessments of emotions during the analysis and examination of documents using several topics were found. In comparison, conventional models concentrated on large rather than minor issues, where accuracy did not seem to be optimized.

Some of the research gap issues found are:

**Research Gap 1**: *Aggregating the different Opinions in a statement.*

A sentence having 2 or more opinion words for some features.

**Research Gap 2**: *Context-dependent opinion words for features.*

High + Price - Negative Opinion, High + Quality - Positive Opinion.

**Research Gap 3**: *Smart ambiguous phrase that can mean anything.*

Common positive or negative opinions related to any company or product.

**Research Gap 4**: *Mentioned right problem but with the wrong emoji.*

Multimodal approaches.

## 7. RESEARCH AGENDA :

1] What approaches and methods are better for classification?
2] What machine learning techniques can improve the performance evaluation of data?
3] What inquire about topics are being tended in Machine learning?
4] What are the main results of the studies before using machine learning?
5] Whatis the level of use of Machine Learning methods?
6] What are the challenges in sentimental analysis when using machine learning techniques?
7] What integrated machine learning techniques can be proposed to give the best performance results?

## 8. ANALYSIS OF RESEARCH AGENDA :

Several methods are available for machine learning, from simple linear methods to extremely non-linear methods, for any particular problem. At first, start for more basic approaches like logistics regression, decision-making boards, and Naive Bayes, then go to SVM, Random Forest, AdaBoost, Naive Networks, etc. and then deeper learning. Compare the accuracy, sensitivity, specificities, and area of the output measurements of the ROC curves of the techniques used for determining the model. Any machine learning techniques that take account of factors such as a matrix of uncertainty, precision, recall, specificity, F1 score, precision-recall or PR curve, ROC curve, can enhance the evaluation of data output. Literary research, which was primarily performed, includes the use of opinion analysis to assess public opinions on health issues and policies and searching for variables that can decide whether or not someone is at risk. Machine Learning has been incorporated into android-based and web-based applications. The obstacle for sentiment analysis is the identification of sarcasm, negation identification, ambiguity of terms, and multipolarity.

## 9. RESEARCH PROPOSAL :

E-commerce is one of the key ways to shop for goods, from basic electronics to expensive objects. While shopping online, customers often rely on reviews from other customers through a credit rating system before finalizing whether to purchase a product. In this way, the seller is motivated by unequal ratings to advertise or show a product of its interest, contrary to the value of its competitors. An online rating system should not only be qualitative and quantitative-based on perception but should also identify conditions under which other users seek to manipulate the system with biased positive and negative ratings that can lead to collusion and exploitation.

It proposes a statistical method to recognize and diagnose unequal ratings, all positive credibility issues, and to diagnose collusion and exploitation. This experimental technique tests the consistency, accuracy, and reminder of the SVM algorithm and logical regression with n-gram performance improvement function selections.



Electronic copy available at: https://ssrn.com/abstract=3674982

## 10. SLOC ANALYSIS OF RESEARCH PROPOSAL :

Although there is continued development in the field of e-commerce, you might easily ask what the future holds. The SLOC review covers strengths, limitations, opportunities, and threats that impact a chosen sector – to decide how ecommerce is prepared for the future.

| | |
|---|---|
| **STRENGTH**<br>1. Gain Competitive Edge by Quicker Response Times<br>2. Better Accuracy<br>3. Useful Tool for Market Research<br>4. Revitalize the Brand | **LIMITATION**<br>1. Completely dependent on dataset<br>2. Less no. of reviews can produce low reputation score.<br>3. Lack of consistent fast internet connection required for processing |
| **OPPURTUNITIES**<br>1. Statistical Analysis System (SAS) can be used for evaluation.<br>2. Various classifier combination can improve accuracy of the result. | **CHALLENGES**<br>1. Aggregating the different Opinions in a statement<br>2. Detecting opinion spam |

## 11. SUGGESTION TO IMPLEMENT RESEARCH PROPOSAL ACTIVITY :

To find unfair ratings, the "all excellent reputation" problem, unequal reputations and collusion and exploitation, the sentiment classification algorithms in any social environments like Facebook, Twitter, etc. Trigrams are used as feature sets for the pre-processing of data and consist of two consecutive terms and may include some contextual information. To introduce and assess the work efficiency, supervised sentiment classification approaches can be used using the R tool.

## 12. CONCLUSION :

This paper summarizes the techniques for machine learning used in the analysis of emotions in the latest periods. Different application areas of sentiment analyzes are also explored in such as industry, politics, public behavior, and finance. Through this paper, the impact of applying data transformations may improvise the achievement of the methods of classification but the type of transformation lieson the dataset and the language it includes. Therefore, look at the details, make a selection of the features, apply transformations and filter the less relevant data making machine learning methods generalize and effective since the computers these days have limits and can't handle them all the data without prior review of any kind. The machine learning methods appear to typically offer the findings are identical and, again, depending on the form of those results. This paper assumes that applications of sentiment analysis will continue to grow in the future and that the implementation of sentiment analytical techniques will be standardized in various systems and services. The proposed future work will focus on three different characteristics chosen to investigate various datasets combining logistic regression and SVM algorithms. It can find unfair positive reviews and unfair negative reviews, reputation issues, and collusion and control through this work. The experimental method can study the accuracy, precision, and recall of both algorithms and can determine accurate and less time feature selection.

## REFERENCES:

[1]Yi, S., & Liu, X. (2020). Machine learning based customer sentiment analysis for recommending shoppers, shops based on customers' review. *Complex & Intelligent Systems*, *1*(1). DOI: https://doi.org/10.1007/s40747-020-00155-2

[2]Vohra, S., & Teraiya, J. (2013). A Comparative Study of Sentiment Analysis Techniques. *International Journal of Information, Knowledge and Research in Computer Engineering*, 2(2),313-317.

[3]Machine Learning & its Applications Outsource to India. (2020). Retrieved on May 18, 2020, from https://www.outsource2india.com/software/articles/machine-learning-applications-how-it-works-who-uses-it.asp



Electronic copy available at: https://ssrn.com/abstract=3674982

[4]Jain, A. P., & Dandannavar, P. (2016). Application of machine learning techniques to sentiment analysis. *Second International Conference on Applied and Theoretical Computing and Communication Technology (ICATccT)*, 1(1). 628–632. DOI: https://doi.org/10.1109/ICATCCT.2016.7912076

[5]Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 5(4), 1093–1113. DOI: https://doi.org/10.1016/j.asej.2014.04.011

[6]Aydogan, E., & Akcayol, M. A. (2016). A comprehensive survey for sentiment analysis tasks using machine learning techniques. 2016 *International Symposium on Innovations in Intelligent Systems and Applications (INISTA*,.1(1) 1–7. DOI: https://doi.org/10.1109/INISTA.2016.7571856

[7]Ahmad, M., Aftab, S., Muhammad, S. S., & Ahmad, S. (2017). Machine learning techniques for sentiment analysis: A review. *International journal of Multi-disciplinary science and Engineering*, 8(3), 27-35.

[8]Mäntylä, M. V., Graziotin, D., & Kuutila, M. (2018). The evolution of sentiment analysis—A review of research topics, venues, and top cited papers. *Computer Science Review*, 27(1), 16–32. DOI: https://doi.org/10.1016/j.cosrev.2017.10.002

[9]Kumar, A., & Sebastian, T. M. (2012). Sentiment Analysis: A Perspective on its Past, Present and Future. *International Journal of Intelligent Systems and Applications*, 4(10), 1–14. DOI: https://doi.org/10.5815/ijisa.2012.10.01

[10] Swathi, R., & Seshadri, R. (2017). Systematic survey on evolution of machine learning for big data. *International Conference on Intelligent Computing and Control* Systems *(ICICCS)*, 1(1), 204–209. DOI: https://doi.org/10.1109/ICCONS.2017.8250711

[11] Kawade, D., & Oza, K. (2017). Sentiment Analysis: Machine Learning Approach. *International Journal of Engineering and Technology*, 9(1), 2183–2186. DOI: https://doi.org/10.21817/ijet/2017/v9i3/1709030151

[12] Patil, H., & Atique, M. (2015). Sentiment Analysis for Social Media: A Survey. *International Conference on Information Science and Security(ICISS)*, 1(1), 1–4. DOI: https://doi.org/10.1109/ICISSEC.2015.7371033

[13] Yang, C., Tang, X., Wong, Y. C., & Wei, C.-P. (2010). Understanding Online Consumer Review Opinions with Sentiment Analysis using Machine Learning. *Pacific Asia Journal of the Association for Information Systems*, 2(1), 73–89. DOI: https://doi.org/10.17705/1pais.02305

[14] Gopu, M., & Swarnalatha, P. (2017). Analyzing customer sentiments using machine learning techniques. *International Journal of Civil Engineering and Technology(IJCIET)*, 8, 1829–1842.

[15] Liu, R., SHI, Y., JI, C., & JIA, M. (2019). A Survey of Sentiment Analysis Based on TransferLearning. *IEEE Access*, 7(1), 85401-85412. DOI:https://doi.org/10.1109/ACCESS.2019.2925059

[16] Naiknaware, B., Kushwaha, B., & Kawathekar, S. (2017). Social Media Sentiment Analysis using Machine Learning Classifiers. *International Journal of Computer Science and Mobile Computing*, 6(6), 465-472.

[17] Singh, J., Singh, G., & Singh, R. (2017). Optimization of sentiment analysis using machine learning classifiers. *Human-Centric Computing and Information Sciences*,7(32), 1-7. DOI: https://doi.org/10.1186/s13673-017-0116-3

[18] Yogi, T. N., & Paudel, N. (2020). Comparative Analysis of Machine Learning Based Classification Algorithms for Sentiment Analysis. *International Journal of Innovative Science, Engineering & Technology*, 7(6),1-9.



Electronic copy available at: https://ssrn.com/abstract=3674982

[19] Mahendran, N., & Mekala, T. (2018). A survey: Sentiment analysis using machine learning techniques for social media analytics. *International Journal of Pure and Applied Mathematics,* 118(8), 419–422.

[20] Patel, A. (2018,). Machine Learning Algorithm Overview. Medium. Retrieved on May 18, 2020 from https://medium.com/ml-research-lab/machine-learning-algorithm-overview-5816a2e6303

[21] Suryawanshi, R., Rajput, A., Kokale, P., & Karve, S. S. (2020). Sentiment Analyzer using Machine Learning. *International Research Journal of Modernization in Engineering Technology and Science*, 02(06),1-12.

[22] K. Sentamilselvan, D. Aneri, A. C. Athithiya, P. Kani Kumar. (2020). Twitter Sentiment Analysis using Machine Learning Techniques. *International Journal of Engineering and Advanced Technology* (IJEAT), 9(3), 1-9. DOI: 0.35940/ijeat.C6281.029320

[23] Godara, N., & Kumar, S. (2020). Twitter Sentiment Classification using Machine Learning Techniques. *Waffen-Und Kostumkunde Journal*, 11(8), 10-20.DOI:010.11205.WJ.2020.V11I7.05.100959

[24] Shirsat, V., Jagdale, R., Shende, K., Deshmukh, S. N., & Kawale, S. (2019). Sentence Level Sentiment Analysis from News Articles and Blogs using Machine Learning Techniques. *International Journal of Computer Sciences and Engineering*, 1(1), 12-17. DOI: https://doi.org/10.26438/ijcse/v7i5.16

[25] Yadav, N., Arora, M. S., & Tech, M. (2020). The Performance of Various Supervised Machine Learning Classification Algorithms in Sentiment Analysis of Online Customer Feedback in Restaurant Sector of Hospitality Industry. *International Journal for Technological Research in Engineering,* 7(11), 1-12.

[26] Gujar, M. A., & Pardeshi, N. G. (2020). Review on A Sentiment Analysis and Predicting Winner for Indian Premier League Using Machine Learning Technique.*International Research Journal of Modernization in Engineering Technology and Science,* 2(6), 963-967.

[27] Arote Rutuja S., Gaikwad Ruchika P., Late Samidha S., Prof. G. B. Gadekar(2020). Online Shopping with Sentimental Analysis for Furniture Shop, *International Research Journal of Modernization in Engineering Technology and Science*,02(05),1-8.

[28] Raju, S. M., & Tarif, A. M. (2020). Real-Time Prediction of BITCOIN Price using Machine Learning Techniques and Public Sentiment Analysis. Retrieved on May 2020 from DOI: http://arxiv.org/abs/2006.14473

[29] Ardianto, R., Rivanie, T., Alkhalifi, Y., Nugraha, F. S., & Gata, W. (2020). Sentiment Analysis on E-Sports for Education Curriculum Using Naive Bayes and Support Vector Machine. *Journal of Computer Sciences and Information*, 13(2), 109–122. DOI: https://doi.org/10.21609/jiki.v13i2.885

[30] Imtiaz, M. N., & Islam, M. K. B. (2020). Identifying Significance of Product Features on Customer Satisfaction Recognizing Public Sentiment Polarity: Analysis of Smart Phone Industry Using Machine-Learning Approaches. *International Journal of Applied Artificial Intelligence*, 1(1), 1–17.DOI: https://doi.org/10.1080/08839514.2020.1787676

[31] Cruz Diaz, N., Taboada, M., & Mitkov, R. (2015). A Machine Learning Approach to Negation and Speculation Detection for Sentiment Analysis. *Journal of the American Society for Information Science and Technology(JASIST)*, 67(9),2118-2136. DOI: https://doi.org/10.1002/asi.23533

[32] Boiy, E., & Moens, M.-F. (2009). A Machine Learning Approach to Sentiment Analysis in Multilingual Web Texts. *International Journal of Information Retrieval Springer*, 12(1), 526–558. DOI: https://doi.org/10.1007/s10791-008-9070-z

[33] Chaturvedi, S., Mishra, V., & Mishra, N. (2017). Sentiment analysis using machine learning for business intelligence. *International Conference on Power, Control, Signals and Instrumentation Engineering IEEE (ICPCSI)*,1(1), 2162–2166. DOI: https://doi.org/10.1109/ICPCSI.2017.8392100



Electronic copy available at: https://ssrn.com/abstract=3674982

[34] Ravi, K., & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. *Knowledge-Based Systems*, 89(1), 14–46. DOI: DOI: https://doi.org/10.1016/j.knosys.2015.06.015

[35] Duan, W., Yu, Y., Cao, Q., & Levy, S. (2015). Exploring the Impact of Social Media on Hotel Service Performance: A Sentimental Analysis Approach. *Cornell Hospitality Quarterly*, 57(3), 282-296 DOI: https://doi.org/10.1177/1938965515620483

[36] Oscar, N., Fox, P. A., Croucher, R., Wernick, R., Keune, J., & Hooker, K. (2017). Machine Learning, Sentiment Analysis, and Tweets: An Examination of Alzheimer's Disease Stigma on Twitter. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, 72(5), 742–751.DOI: https://doi.org/10.1093/geronb/gbx014

[37] Moh, M., Gajjala, A., Gangireddy, S. C. R., & Moh, T.-S. (2015). On Multi-tier Sentiment Analysis Using Supervised Machine Learning. *International Conference on Web Intelligence and Intelligent Agent Technology* IEEE/WIC/ACM (WI-IAT), 1(1), 341–344. DOI: https://doi.org/10.1109/WI-IAT.2015.154

[38] Jabbar, J., Urooj, I., JunSheng, W., & Azeem, N. (2019). Real-time Sentiment Analysis on E-Commerce Application. *International Conference on Networking, Sensing and Control(ICNSC)*,1(1),391–396. DOI: https://doi.org/10.1109/ICNSC.2019.8743331

[39] Tuhin, R. A., Paul, B., Nawrine, F., Akter, M., & Das, A. (2019). An Automated System of Sentiment Analysis from Bangla Text using Supervised Learning Techniques. *International Conference on Computer and Communication Systems (ICCCS)*,1(1), 360–364. DOI: https://doi.org/10.1109/CCOMS.2019.8821658

[40] Li, N., & Wu, D. D. (2010). Using text mining and sentiment analysis for online forums hotspot detection and forecast. *Decision Support Systems Elsevier*, 48(2), 354–368. DOI: https://doi.org/10.1016/j.dss.2009.09.003

[41] Tripathy, A., Agrawal, A., & Rath, S. (2016). Classification of Sentiment Reviews using N-gram Machine Learning Approach. *Expert Systems with Applications*, 57(1),117-126 DOI: https://doi.org/10.1016/j.eswa.2016.03.028

[42] Gamal, D., Alfonse, M., El-Horbarty, E.-S., & M.Salem, A.-B. (2019). An Evaluation of Sentiment Analysis on Smart Entertainment and Devices Reviews. *International JournalInformation Theories and Applications*, 26 (2), 147–164.

[43] Kauffmann, E., Peral, J., Gil, Ferrández, A., Sellers-Rubio, R., & Mora, H. (2019). Managing Marketing Decision-Making with Sentiment Analysis: An Evaluation of the Main Product Features Using Text Data Mining. Sustainability, *Digital Marketing for Sustainable Growth: Business Models and Online Campaign*,11(15), 1-17. DOI:https://doi.org/10.3390/su11154235

[44] Kakulapati, V. (2017). A Panoptics of Sentimental Analysis. *International Journal of Advanced Research in Computer Science*, 8(1), 1036–1041.DOI: https://doi.org/10.26483/ijarcs.v8i7.4448

[45] Aziz, A., Starkey, A., & Madi, E. (2019). Predicting Supervise Machine Learning Performances for sentiment analysis using contextual-based approaches. *IEEE Acces*, 8(1), 17722-17733. DOI: https://doi.org/10.1109/ACCESS.2019.2958702

[46] Abo, M., Shah, N., Balakrishnan, V., & Abdelaziz, A. (2018). Sentiment analysis algorithms: Evaluation performance of the Arabic and English language. *IEEE Expert*,1(1), 1-5. DOI: https://doi.org/10.1109/ICCCEEE.2018.8515844

[47] Trivedi, S. (2017, April 1). Sentiment analysis of Indian movie review with various feature selection techniques. *International Conference on Advances in Computer Applications (ICACA)*, 1(1), 181-185. DOI: https://doi.org/10.1109/ICACA.2016.7887947



Electronic copy available at: https://ssrn.com/abstract=3674982

[48] Bansal, A., Gupta, C. L., & Muralidhar, A. (2019). A Sentimental Analysis for YouTube Data using Supervised Learning Approach. *International Journal of Engineering and Advanced Technology (IJEAT)*, 8 (5).1-12.

[49] Alsaeedi, A., & Khan, M. (2019). A Study on Sentiment Analysis Techniques of Twitter Data. *International Journal of Advanced Computer Science and Applications*, 10(1), 361–374. DOI: https://doi.org/10.14569/IJACSA.2019.0100248

[50] Kawade, D., & Oza, K. (2017). Sentiment Analysis: Machine Learning Approach. *International Journal of Engineering and Technology*, 9(1), 2183–2186. DOI: https://doi.org/10.21817/ijet/2017/v9i3/1709030151

[51] S, M., & Muthu, R. (2017, June 23). Sentiment Analysis on Speaker Specific Speech Data. *International Conference on Intelligent Computing and Control (I2C2)*, 1(1), 1-5, DOI: https://doi.org/10.1109/I2C2.2017.8321795

[52] Sadhasivam, J., & Babu, R. (2019). Sentiment Analysis of Amazon Products Using Ensemble Machine Learning Algorithm. *International Journal of Mathematical, Engineering and Management Sciences*, 4(1), 508–520. DOI: https://doi.org/10.33889/IJMEMS.2019.4.2-041

[53] Hasan, A., Moin, S., Karim, A., & Shamshirband, D. (2018). Machine Learning-Based Sentimental Analysis for Twitter Accounts. *Mathematical and Computational Applications*, 23(1), 11-32. DOI: https://doi.org/10.3390/mca23010011

[54] Boiy, E., & Moens, M.-F. (2009). A machine learning approach to sentiment analysis in multilingual Web texts. *Information Retrieval Springer*. 12(5), 526–558. DOI: https://doi.org/10.1007/s10791-008-9070-z

[55] Sultana, N., Kumar, P., Patra, M., Chandra, S., & Alam, S. (2019). Sentiment Analysis for Product Review. *International Journal of Soft Computing*, 09(1), 7-28. DOI: https://doi.org/10.21917/ijsc.2019.0266

[56] Axhiu, M., Veljanoska, F., Ciglovska, B., & Husejni, M. (2014). The Usage of Sentiment Analysis for Hearing the Voice of the Customer and Improving Businesses. *Journal of Educational and Social Research*, 4(4),1-17. DOI: https://doi.org/10.5901/jesr.2014.v4n4p401

[57] S, N., M, A., & K, A. (2017). Survey on Sentiment Analysis of Stock Market, *International Journal of Research Granthaaayah*,5(4), 69–75. DOI: https://doi.org/10.5281/zenodo.572298

[58] Raza, H., Faizan, M., Hamza, A., Mushtaq, A., & Akhtar, N. (2019). Scientific Text Sentiment Analysis using Machine Learning Techniques. *International Journal of Advanced Computer Science and Applications*,10(12), 157-165. DOI: https://doi.org/10.14569/IJACSA.2019.0101222

[59] Isah, H., Neagu, D., & Trundle, P. (2015). Social Media Analysis for Product Safety using Text Mining and Sentiment Analysis. *Workshop on Computational Intelligence, UKCI – Proceedings*,1(1), 1-7. DOI: https://doi.org/10.1109/UKCI.2014.6930158

[60] Cao, J., Ke, Z., Wang, H., Cheng, J., Qiao, F., Wen, D., & Gao, Y. (2014). Web-Based Traffic Sentiment Analysis: Methods and Applications. Intelligent Transportation Systems, *IEEE Transactions*, 15(1), 844–853. DOI: https://doi.org/10.1109/TITS.2013.2291241

[61] Duric, A., & Song, F. (2012). Song, F.: Feature Selection for Sentiment Analysis based on Content and Syntax Models. Decision Support Systems 53, 704-711. *Decision Support Systems*, 53(1), 704–711. DOI: https://doi.org/10.1016/j.dss.2012.05.023

[62] Go, A., Bhayani, R., & Huang, L. (2009). Twitter sentiment classification using distant supervision. Processing, *CS224N project report Stanford*, 1(12), 2009-2016.



Electronic copy available at: https://ssrn.com/abstract=3674982

[63] Buczak, A., & Guven, E. (2015). A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. *IEEE Communications Surveys & Tutorials*, 18(2),1153-1176. DOI: https://doi.org/10.1109/COMST.2015.2494502

[64] Das, S., Behera, R., Kumar, M., & Rath, S. (2018). Real-Time Sentiment Analysis of Twitter Streaming data for Stock Prediction. *Procedia Computer Science*, 132(1), 956–964. DOI: https://doi.org/10.1016/j.procs.2018.05.111

[65] Asghar, Dr. M. (2014). Detection and Scoring of Internet Slangs for Sentiment Analysis Using SentiWordNet. *Life Science Journal*, 11(1), 66–72. DOI: https://doi.org/10.6084/M9.FIGSHARE.1609621

[66] Engonopoulos, N., Lazaridou, A., Paliouras, G., & Chandrinos, K. (2011). ELS: A word-level method for entity-level sentiment analysis. *Proceedings of the International Conference on Web Intelligence, Mining and Semantics*,12(1),1-9. DOI: https://doi.org/10.1145/1988688.1988703

[67] Ravi, K., Vasili, S., Vadlamani, R., & Mohan, L. (2015). Sentiment analysis applied to Educational Sector. *International Conference on Computational Intelligence and Computing Research (ICCIC)*, 1(1), 1–6. DOI: https://doi.org/10.1109/ICCIC.2015.7435667

[68] Nopp, C., & Hanbury, A. (2015). Detecting Risks in the Banking System by Sentiment Analysis. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*,1(1), 591–600. DOI: https://doi.org/10.18653/v1/D15-1071

[69] Hassan, A. U., Hussain, J., Hussain, M., Sadiq, M., & Lee, S. (2017, October 18). Sentiment Analysis of Social Networking Sites (SNS) Data using Machine Learning Approach for the Measurement of Depression. *International Conference on Information and Communication Technology Convergence (ICTC*), 1(1), 138-140. DOI: https://doi.org/10.1109/ICTC.2017.8190959

[70] Ortigosa, A., Martín, J. M., & Carro, R. M. (2014). Sentiment analysis in Facebook and its application to e-learning. *Computers in Human Behavior*, 31(1), 527–541. DOI: https://doi.org/10.1016/j.chb.2013.05.024

[71] Pang, B., & Lee, L. (2004). A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, 1(1), 271-278. DOI: https://doi.org/10.3115/1218955.1218990

[72] Li, N., Liang, X., Li, X., Wang, C., & Wu, D. D. (2009). Network Environment and Financial Risk Using Machine Learning and Sentiment Analysis. *International Journal of Human and Ecological Risk Assessment*,15(2), 227–252. DOI: https://doi.org/10.1080/10807030902761056

[73] Laryea, B. N. L., Choi, C.-H., Jung, I.-S., Lee, K.-H., & Cho, W.-S. (2015). Web application for sentiment analysis using supervised machine learning. *International Journal of Software Engineering and Its Applications*, 9(1), 191–200.DOI: https://doi.org/10.14257/ijseia.2015.9.1.17

[74] Singh, V. K., Piryani, R., Uddin, A., Waila, P., & Marisha. (2013). Sentiment analysis of textual reviews; Evaluating machine learning, unsupervised and SentiWordNet approaches. *International Conference on Knowledge and Smart Technology (KST)*, 1(1), 122–127. DOI: https://doi.org/10.1109/KST.2013.6512800

[75] Ghiassi, M., & Lee, S. (2018). A domain transferable lexicon set for Twitter sentiment analysis using a supervised machine learning approach. *Expert Systems with Applications*, 106(1), 197–216. DOI: https://doi.org/10.1016/j.eswa.2018.04.006

[76] Samal, B., Behera, A. K., & Panda, M. (2017). Performance analysis of supervised machine learning techniques for sentiment analysis. Third*International Conference on Sensing, Signal Processing and Security (ICSSS*), 1(1), 128–133. DOI: https://doi.org/10.1109/SSPS.2017.8071579



Electronic copy available at: https://ssrn.com/abstract=3674982

[77] Valencia, F., Gómez-Espinosa, A., & Valdés-Aguirre, B. (2019). Price Movement Prediction of Cryptocurrencies Using Sentiment Analysis and Machine Learning. *International and interdisciplinary Journal Entropy*, 21(6), 589. DOI: https://doi.org/10.3390/e21060589

[78] Nair, D. S., Jayan, J. P., Rajeev R.R, & Sherly, E. (2015). Sentiment Analysis of Malayalam film review using machine learning techniques. *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 1(1), 2381–2384. DOI: *https://doi.org/10.1109/ICACCI.2015.7275974*

[79] Daeli, N. O. F., & Adiwijaya, A. (2020). Sentiment Analysis on Movie Reviews using Information Gain and K-Nearest Neighbor. *Journal of Data Science and Its Applications*, 3(1), 1–7.DOI: https://doi.org/10.34818/jdsa.2020.3.22

[80] Altrabsheh, N., Cocea, M., & Fallahkhair, S. (2014). Sentiment Analysis: Towards a Tool for Analysing Real-Time Students Feedback. *IEEE International Conference on Tools with Artificial Intelligence*, 1(1), 419–423. DOI: https://doi.org/10.1109/ICTAI.2014.70

[81] Dadvar, M., Hauff, C., & de Jong, F. (2011). Scope of negation detection in sentiment analysis. *Dutch- Belgian Information Retrieval Workshop*, 1(1),16-20.

[82] Kumar, S., Gahalawat, M., Roy, P. P., Dogra, D. P., & Kim, B.-G. (2020). Exploring Impact of Age and Gender on Sentiment Analysis Using Machine Learning. International Journal of Electronics, 9(2), 374. DOI: https://doi.org/10.3390/electronics9020374

[83] Wu, T., Weld, D. S., & Heer, J. (2019). Local Decision Pitfalls in Interactive Machine Learning: An Investigation into Feature Selection in Sentiment Analysis. *ACM Transactions on Computer-Human Interaction*, 26(4), 24–27.DOI: https://doi.org/10.1145/3319616

[84] Waila, P., Marisha, Singh, V. K., & Singh, M. K. (2012). Evaluating Machine Learning and Unsupervised Semantic Orientation approaches for sentiment analysis of textual reviews. *IEEE International Conference on Computational Intelligence and Computing Research*, 1(1), 1–6. DOI: https://doi.org/10.1109/ICCIC.2012.6510235

[85] Ashok, M., Rajanna, S., Joshi, P. V., & Sowmya Kamath S. (2016). A personalized recommender system using Machine Learning based Sentiment Analysis over social data. *IEEE Conference on Electrical, Electronics and Computer Science (SCEECS)*, 1(1), 1–6. DOI: https://doi.org/10.1109/SCEECS.2016.7509354

[86] Ghosh, M., & Sanyal, G. (2018). An ensemble approach to stabilize the features for multi-domain sentiment analysis using supervised machine learning. *International Journal of Big Data*, 5(1), 44-49. DOI: https://doi.org/10.1186/s40537-018-0152-5

[87] Liu, H., & Cocea, M. (2017). Fuzzy rule based systems for interpretable sentiment analysis. *International Conference on Advanced Computational Intelligence (ICACI)*,1(1), 129–136. DOI: https://doi.org/10.1109/ICACI.2017.7974497

[88] Hew, K. F., Hu, X., Qiao, C., & Tang, Y. (2020). What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach. *Computers & Education Elsevie*, 145(1), 1-17.DOI:https://doi.org/10.1016/j.compedu.2019.103724

[89] Shi, H.-X., & Li, X.-J. (2011). A sentiment analysis model for hotel reviews based on supervised learning. *International Conference on Machine Learning and Cybernetics*, 3(1), 950–954. DOI: https://doi.org/10.1109/ICMLC.2011.6016866

[90] Troussas, C., Virvou, M., Espinosa, K. J., Llaguno, K., & Caro, J. (2013). Sentiment analysis of Facebook statuses using Naive Bayes classifier for language learning. *International Conference on Information, Intelligence, Systems and Applications (IISA)*, 1(1), 1–6. DOI: https://doi.org/10.1109/IISA.2013.6623713



Electronic copy available at: https://ssrn.com/abstract=3674982

[91] Hausler, J., Ruscheinsky, J., & Lang, M. (2018). News-based sentiment analysis in real estate: A machine learning approach. *Journal of Property Research,* 35(4), 344–371. DOI: https://doi.org/10.1080/09599916.2018.1551923

[92] Dey, L., Chakraborty, S., Biswas, A., Bose, B., & Tiwari, S. (2016). Sentiment Analysis of Review Datasets Using Naïve Bayes 'and K-NN Classifier. *International Journal of Information Engineering and Electronic Business*, 8(1), 54–62.DOI: https://doi.org/10.5815/ijieeb.2016.04.07

[93] Sivakumar, M., & Reddy, U. S. (2017). Aspect based sentiment analysis of student's opinion using machine learning techniques. *International Conference on Inventive Computing and Informatics (ICICI).* 1(1), 726–731.DOI: https://doi.org/10.1109/ICICI.2017.8365231

[94] Esparza, G. G., de-Luna, A., Zezzatti, A. O., Hernandez, A., Ponce, J., Álvarez, M., Cossio, E., & de Jesus Nava, J. (2018). A Sentiment Analysis Model to Analyze Students Reviews of Teacher Performance Using Support Vector Machines. *International Symposium on Distributed Computing and Artificial Intelligence*,620 (1), 1-7. DOI: https://doi.org/10.1007/978-3-319-62410-5_19

[95] Das Adhikari, N., Alka, A., Kushwaha, J., & Nayak, A. (2018, July 14). Sentiment Classifier and Analysis for Epidemic Prediction.*International Conference on Computer Science & Information Technology (CSCP).* 1(1), 31-48. DOI:https://doi.org/10.5121/csit.2018.81004

[96] Gurkhe, D., Pal, N., & Bhatia, R. (2014). Effective Sentiment Analysis of Social Media Datasets using Naive Bayesian Classification. *International Journal of Computer Applications*. 99(1|),1–4.DOI: https://doi.org/10.5120/17430-8274

[97] Rohini, V., Thomas, M., & Latha, C. A. (2016). Domain based sentiment analysis in regional Language-Kannada using machine learning algorithm. *IEEE International Conference on Recent Trends in Electronics, Information Communication Technology (RTEICT).* 1(1), 503–507. DOI: https://doi.org/10.1109/RTEICT.2016.7807872

[98] Shuhidan, S. M., Hamidi, S. R., Kazemian, S., Shuhidan, S. M., & Ismail, M. A. (2018). Sentiment Analysis for Financial News Headlines using Machine Learning Algorithm. *International Conference on Kansei Engineering and Emotion Research,*739(1), (64–72). DOI: https://doi.org/10.1007/978-981-10-8612-0_8

[99] Maharani, W. (2013). Microblogging sentiment analysis with lexical based and machine learning approaches. *International Conference of Information and Communication Technology (ICoICT),* 1(1), 439–443. DOI: https://doi.org/10.1109/ICoICT.2013.6574616

[100] Sentiment Analysis: Types, Tools, and Use Cases. (2018). AltexSoft. Retrieved on June 23, 2020, from https://www.altexsoft.com/blog/business/sentiment-analysis-types-tools-and-use-cases.

**********



Electronic copy available at: https://ssrn.com/abstract=3674982